

**Sławomir Breiter**

*Matematyczne podstawy mechaniki nieba*

I r. Astron. (II st.)

wersja 12.10.2022

# Rozdział 1

## Równania ruchu w formalizmie newtonowskim

Skoro mechanika nieba jest nauką o ruchu ciał niebieskich, to umiejętność formułowania równań ruchu ma znaczenie podstawowe. Zaczniemy od przypomnienia klasycznego formalizmu Newtona, który nakłada najmniej ograniczeń na postać sił działających w rozpatrywanym układzie. Skupimy się na problemach dynamiki punktów materialnych, pozostawiając na uboczu kwestię ruchu obrotowego bryły sztywnej.

### 1.1 Formalizm newtonowski

Najbardziej klasyczne podejście do mechaniki korzysta z równań ruchu w postaci wprowadzonej przez Newtona w trzech zasadach dynamiki. Przez formalizm newtonowski rozumiemy rozpatrywanie ruchu w kategoriach takich pojęć podstawowych jak układ inercjalny (I zasada) oraz siła i przyspieszenie (II zasada) opisywane w „naturalnych” zmiennych kartezjańskich. Oczywiście, można w ramach formalizmu newtonowskiego wprowadzać inne układy zmiennych, ale przejście do nich wymaga najpierw ich zdefiniowania jako funkcji kartezjańskich położenia i prędkości.

Jeśli więc mamy układ  $N$  punktów materialnych, to – w świetle drugiej zasady dynamiki – jego równania ruchu mają postać

$$m_i \ddot{\mathbf{r}}_i = \mathbf{F}_i(\mathbf{r}_1, \dots, \mathbf{r}_N, \dot{\mathbf{r}}_1, \dots, \dot{\mathbf{r}}_N, t), \quad i = 1, \dots, N, \quad (1.1)$$

gdzie siła  $\mathbf{F}_i$  może zależeć od położenia  $\mathbf{r}_i$  i prędkości  $\dot{\mathbf{r}}_i$  wszystkich ciał oraz jawnie od czasu i – co najważniejsze – nie musi spełniać żadnych dodatkowych założeń. W newtonowskiej mechanice klasycznej siła nie zależy od

pochodnych  $\mathbf{r}_i$  rzędu wyższego niż pierwszy. Przykładem zastosowania formalizmu newtonowskiego mogą być znane z *Wstępu do mechaniki nieba* równania ruchu zagadnienia dwóch ciał i przeprowadzone następnie całkowanie tego zagadnienia.

Stosując pojęcie pędu  $i$ -tej cząstki  $\mathbf{p}_i$ , możemy zapisać równania (1.1) w postaci

$$\begin{aligned}\dot{\mathbf{r}}_i &= \frac{\mathbf{p}_i}{m_i}, \\ \dot{\mathbf{p}}_i &= \mathbf{F}_i(\mathbf{r}_1, \dots, \mathbf{r}_N, \mathbf{p}_1/m_1, \dots, \mathbf{p}_N/m_N, t),\end{aligned}\quad (1.2)$$

gdzie pierwsze równanie pośrednio definiuje pęd jako iloczyn masy i prędkości  $\mathbf{p}_i = m_i \dot{\mathbf{r}}_i$ , zaś drugie ustala równość między prędkością zmian pędu a działającą siłą. Jedną z zalet układu (1.2) jest to, że daje on przedsmak równań kanonicznych. W ramach formalizmu Newtona ważniejsze jednak jest to, że w ten sposób sprowadzamy równania ruchu do standardowej postaci z teorii równań różniczkowych, otrzymując zamiast  $3N$  równań drugiego rzędu – układ  $6N$  równań rzędu pierwszego

$$\dot{\mathbf{w}} = \mathbf{F}(\mathbf{w}, t), \quad (1.3)$$

gdzie **wektor stanu** ma składowe

$$\mathbf{w} = \begin{pmatrix} \mathbf{r}_1 \\ \dots \\ \mathbf{r}_N \\ \mathbf{p}_1 \\ \dots \\ \mathbf{p}_N \end{pmatrix} \equiv \text{col}(\mathbf{r}_1, \dots, \mathbf{r}_N, \mathbf{p}_1, \dots, \mathbf{p}_N),$$

a prawe strony równań ruchu to

$$\mathbf{F} = \text{col}\left(\frac{\mathbf{p}_1}{m_1}, \dots, \frac{\mathbf{p}_N}{m_N}, \mathbf{F}_1(\mathbf{w}, t), \dots, \mathbf{F}_N(\mathbf{w}, t)\right).$$

Często jednak wygodniej jest posługiwać się prędkościami  $\mathbf{v}_i$  zamiast pędów. Używamy wtedy zamiast (1.3) układu

$$\dot{\boldsymbol{\rho}} = \mathbf{G}(\boldsymbol{\rho}, t), \quad (1.4)$$

gdzie

$$\boldsymbol{\rho} = \text{col}(\mathbf{r}_1, \dots, \mathbf{r}_N, \mathbf{v}_1, \dots, \mathbf{v}_N),$$

oraz

$$\mathbf{G} = \text{col} \left( \mathbf{v}_1, \dots, \mathbf{v}_N, \frac{\mathbf{F}_1(\boldsymbol{\rho}, t)}{m_1}, \dots, \frac{\mathbf{F}_N(\boldsymbol{\rho}, t)}{m_N} \right).$$

Wbrew utartym schematom, układ z jawną zależnością siły od czasu jest tylko **szczególnym przypadkiem** układów zachowawczych (bez jawnej zależności od czasu). Wystarczy wprowadzić **rozszerzony wektor stanu** o  $6N + 1$  składowych  $\mathbf{w}^* = \text{col}(u, \mathbf{w})$ , z dodatkową zmienną  $u$  oraz wektor prawych stron  $\mathbf{F}^* = \text{col}(1, \mathbf{F})$  i mamy już równania zachowawcze

$$\dot{\mathbf{w}}^* = \mathbf{F}^*(\mathbf{w}^*), \quad (1.5)$$

z pierwszą składową

$$\dot{u} = 1,$$

czyli  $u = t + \text{const}$ . W prawych stronach równań nie występuje już czas  $t$  (zmienna niezależna układu), gdyż jego miejsce zajął „sobowtór”  $u$ , będący formalnie jedną ze zmiennych zależnych  $\boldsymbol{\rho}$ .

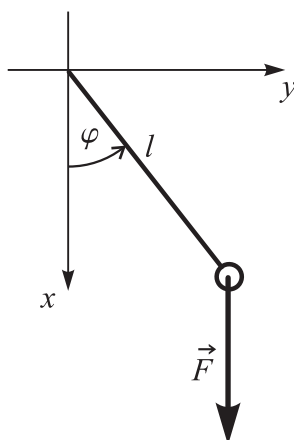
Jeśli równania ruchu tworzą układ równań różniczkowych rzędu  $2M$  bez jawnej zależności od czasu, to mówimy, że układ ma  $M$  **stopni swobody**. Stopniem swobody nazwiemy każda parę współrzędnych położenie-pęd (lub położenie-prędkość), wymaganą do opisu ruchu. Jeśli rząd układu jest nieparzysty, a wiąże się to z usuwaniem jawnej zależności od czasu, mówimy żargonowo o połowie stopnia swobody (zmienna  $u$  nie ma stowarzyszonego pędu jako niezależnej zmiennej, gdyż jej prędkość jest z definicji stała i wynosi 1 niezależnie od warunków początkowych). I tak, układ opisany równaniami (1.5) posiada „ $3N$  i pół” stopnia swobody. Także w sytuacjach, gdy utrzymaliśmy jawną zależność od czasu i formalnie posługujemy się nadal układem rzędu  $2M$ , dorzucamy pół stopnia swobody, gdyż możliwe jest rozszerzenie wektora stanu do wymiaru  $2M + 1$ , nawet jeśli go nie wykonaliśmy.

Wspomniana wyżej dowolność siły sprawia, że formalizm newtonowski jest najogólniejszym sposobem analizy ruchu, dającym się zastosować we wszystkich zagadnieniach mechaniki klasycznej. Niestety, zaleta maksymalnej ogólności oznacza, że nie sposób powiedzieć niczego konkretnego o rozwiązaniach dowolnych równań ruchu w postaci Newtona, gdyż niewiele można powiedzieć o wszystkich możliwych zagadnieniach. Inną wadą jest żmudność przejścia od kartezjańskich położenia i prędkości do bardziej dogodnych zmiennych; zilustrujemy to przykładem równań ruchu wahadła mate-

matycznego.

## 1.2 Wahadło matematyczne, oscylator harmoniczny i oscylator Duffinga

Rozpatrzmy równania ruchu Newtona dla płaskiego wahadła matematycznego o stałej masie  $m$ , zawieszonoego na nieważkim pręcie o długości  $l$  w polu grawitacyjnym o przyspieszeniu  $g$  (Rys. 1.1). Co prawda łatwo jest napisać,



Rysunek 1.1: Wahadło matematyczne.

że dla  $\mathbf{r} = (x, y)^T$  oraz  $\mathbf{F} = (mg, 0)^T$  równania ruchu przyjmują postać

$$\begin{aligned}\ddot{x} &= g, \\ \ddot{y} &= 0,\end{aligned}\tag{1.6}$$

ale równania te są tylko punktem wyjścia do dalszej analizy, wymagającej uwzględnienia równania więzów

$$x^2 + y^2 = l^2.\tag{1.7}$$

Dopiero równanie (1.7) zawiera istotną wiadomość, że układ ma nie dwa, a tylko jeden stopień swobody. Ponieważ wahadło, jako formalny model matematyczny, pełni istotną rolę w mechanice nieba, poprowadźmy dalsze wyprowadzenie wprowadzając jako zmienną kąt  $\varphi$ , definiowany poprzez

$$x = l \cos \varphi, \quad y = l \sin \varphi.\tag{1.8}$$

Różniczkując równania (1.8) dochodzimy do

$$\dot{x} = -l\dot{\varphi} \sin \varphi, \quad \dot{y} = l\dot{\varphi} \cos \varphi,$$

i dalej

$$\begin{aligned} \ddot{x} &= -l\dot{\varphi}^2 \cos \varphi - l\ddot{\varphi} \sin \varphi, \\ \ddot{y} &= -l\dot{\varphi}^2 \sin \varphi + l\ddot{\varphi} \cos \varphi. \end{aligned}$$

Porównanie ich z równaniami (1.6) prowadzi do

$$-l\dot{\varphi}^2 \cos \varphi - l\ddot{\varphi} \sin \varphi = g, \quad (1.9)$$

$$-l\dot{\varphi}^2 \sin \varphi + l\ddot{\varphi} \cos \varphi = 0. \quad (1.10)$$

Należy teraz pomnożyć obie strony (1.9) przez  $-l^{-1} \sin \varphi$  zaś (1.10) przez  $l^{-1} \cos \varphi$  i dopiero teraz, dodając stronami, uzyskamy

$$\ddot{\varphi} = -\omega_0^2 \sin \varphi, \quad (1.11)$$

gdzie

$$\omega_0 = \sqrt{\frac{g}{l}}.$$

Równanie różniczkowe drugiego rzędu (1.11) można także przedstawić jako układ dwóch równań pierwszego rzędu: jeśli wprowadzimy prędkość kątową  $\Phi$ , to mamy

$$\begin{aligned} \dot{\varphi} &= \Phi, \\ \dot{\Phi} &= -\omega_0^2 \sin \varphi, \end{aligned} \quad (1.12)$$

czyli  $\dot{\boldsymbol{\rho}} = \mathbf{G}(\boldsymbol{\rho})$ , gdzie

$$\boldsymbol{\rho} = (\rho_1, \rho_2)^T = (\varphi, \Phi)^T,$$

oraz

$$\mathbf{G} = (\rho_2, -\omega_0^2 \sin \rho_1)^T = (\Phi, -\omega_0^2 \sin \varphi)^T.$$

W przedstawionym tu przykładzie sytuację komplikowało pojawienie się więzów. Można wysunąć zastrzeżenie, że przecież w zagadnieniach ruchu ciał niebieskich nie pojawiają się więzy – trudno wyobrazić sobie planetę na sznurku. A jednak więzy mogą się zjawić nie tylko w postaci fizycznej. Każda całka ruchu

$$\Psi(\boldsymbol{\rho}, t) = \text{const},$$

definiuje w istocie powierzchnię, na której ruch musi się odbywać, pełniąc rolę „matematycznego sznurka”. Wykorzystanie całek ruchu do obniżenia

liczby stopni swobody jest w formalizmie Newtona bardziej zawile niż w ramach formalizmu Lagrange'a czy Hamiltona.

Założmy teraz, że maksymalne wychylenie wahadła (w radianach) jest niewielkie. Rozwijając prawą stronę równania (1.11) w szereg potęgowy wzorem MacLaurina, mamy

$$\sin \varphi = \varphi - \frac{\varphi^3}{6} + O(\varphi^5).$$

Zaniedbując wyrazy rzędu  $\varphi^5$  (lub mniejsze), sprowadzimy równania wahadła do postaci

$$\begin{aligned} \dot{\varphi} &= \Phi, \\ \dot{\Phi} &= -\omega_0^2 \varphi + \frac{1}{6} \omega_0^2 \varphi^3, \end{aligned} \quad (1.13)$$

czyli

$$\ddot{\varphi} = -\omega_0^2 \varphi + \frac{1}{6} \omega_0^2 \varphi^3.$$

Układ ten nazywamy **oscylatorem Duffinga**.

Najdalej idące uproszczenie polega na przyjęciu  $\sin \varphi \approx \varphi$ , z błędem rzędu  $O(\varphi^3)$ . Równania ruchu wahadła sprowadzają się wtedy do równań oscylatora harmonicznego

$$\begin{aligned} \dot{\varphi} &= \Phi, \\ \dot{\Phi} &= -\omega_0^2 \varphi, \end{aligned} \quad (1.14)$$

czyli

$$\ddot{\varphi} = -\omega_0^2 \varphi.$$

Układ ten posiada proste rozwiązanie ogólne

$$\begin{aligned} \varphi(t) &= A \sin(\omega_0 t + \chi), \\ \Phi(t) &= A \omega_0 \cos(\omega_0 t + \chi), \end{aligned} \quad (1.15)$$

zależne od dwóch stałych dowolnych: amplitudy  $A$  i fazy początkowej  $\chi$ . Jest to w istocie rodzina rozwiązań – elips na płaszczyźnie fazowej  $(\varphi, \Phi)$ . Rozwiązanie szczególne, czyli pojedynczą trajektorię na płaszczyźnie fazowej, otrzymamy zadając warunki początkowe w epoce  $t_0$

$$\varphi_0 = \varphi(t_0), \quad \Phi_0 = \Phi(t_0),$$

i wyliczając z nich stałe dowolne  $A, \chi$ .

Zarówno wahadło jak i oscylator Duffinga także posiadają rozwiązania ścisłe, zależne od dwóch stałych dowolnych, ale wymagają one wprowadzenia tak zwanych funkcji eliptycznych.

### 1.3 Metoda uzmienniania stałych

Pozostając nadal na poziomie ogólności formalizmu newtonowskiego, wprowadzimy teraz nowy, dość szczególny rodzaj zmiennych.

Z podstawowego kursu analizy matematycznej<sup>1</sup> można wynieść przekonanie, że **metoda uzmienniania stałych** jest metodą *rozwiązywania* równań różniczkowych zwyczajnych. Jest to prawdziwe, jeśli chodzi o równania liniowe niejednorodne, lecz w ogólnym przypadku równań nieliniowych metoda uzmienniania stałych jest jedynie sposobem *przekształcenia* równań różniczkowych do postaci, która czasem może uprościć dalsze kroki zmierzające do ich rozwiązania.

Aby nie wdawać się w zbędne uogólnienia, skoncentrujemy się na przypadku układu mechanicznego o  $M$  stopniach swobody. Oznaczając przez  $\mathbf{r}, \mathbf{v} \in \mathbb{R}^M$  położenie i prędkość układu, możemy podać jego równania ruchu

$$\begin{aligned}\dot{\mathbf{r}} &= \mathbf{v}, \\ \dot{\mathbf{v}} &= \mathbf{F}_0(\mathbf{r}, \mathbf{v}).\end{aligned}\tag{1.16}$$

Założmy, że układ z siłą (na jednostkę masy)  $\mathbf{F}_0$  jest całkowny i potrafimy znaleźć jego rozwiązanie

$$\mathbf{r} = \mathbf{r}(\mathbf{C}, t), \quad \mathbf{v} = \mathbf{v}(\mathbf{C}, t),\tag{1.17}$$

jawnie zależne od czasu  $t$  oraz od stałych dowolnych  $\mathbf{C} \in \mathbb{R}^{2M}$  wyznaczanych z warunków początkowych. Zauważmy, że po lewej stronie równości mamy wielkości fizyczne: położenie  $\mathbf{r}$  i prędkość  $\mathbf{v}$ , natomiast po prawej – **funkcje**  $\mathbf{r}(\mathbf{C}, t)$ ,  $\mathbf{v}(\mathbf{C}, t)$  czasu i stałych, które opisują zmiany położenia i prędkości. Zagadnienie (1.16) możemy nazwać **zagadnieniem definiującym**, a jego rozwiązanie (1.17) – **rozwiązaniem definiującym**.

Wprowadźmy teraz do układu (1.16) dodatkową siłę  $\mathbf{P}$

$$\begin{aligned}\dot{\mathbf{r}} &= \mathbf{v}, \\ \dot{\mathbf{v}} &= \mathbf{F}_0(\mathbf{r}, \mathbf{v}) + \mathbf{P}(\mathbf{r}, \mathbf{v}).\end{aligned}\tag{1.18}$$

Istotą metody uzmienniania stałych jest przyjęcie, że mimo pojawienia się „zaburzenia”  $\mathbf{P}$ , będziemy nadal używać wzorów (1.17) wywodzących się z zagadnienia definiującego. Ponieważ utrzymujemy postać rozwiązania jako

<sup>1</sup>np. A. Sołtysiak *Analiza matematyczna* rozdz. 9.



funkcji stałych dowolnych i czasu, musimy dopuścić możliwość, że dotychczasowe stałe dowolne  $\mathbf{C}$  przestają być stałymi i przybierają postać nieznanych funkcji czasu  $\mathbf{C}(t)$ . W tej sytuacji, czeka nas wyprowadzenie równań różniczkowych opisujących ewolucję **uzmiennionych stałych**  $\mathbf{C}(t)$ , nazywanych także **stałymi oskulacyjnymi** (lub zmiennymi oskulacyjnymi).

Ogólna procedura zastąpienia równań (1.18) układem  $2M$  równań dla  $\dot{\mathbf{C}}$  wykorzystuje nawiasy Lagrange'a i można ją znaleźć w starszych podręcznikach mechaniki klasycznej lub mechaniki nieba. W praktyce jednak, wygodniej jest analizować każdy przypadek osobno i wykorzystać w maksymalnym stopniu jego specyfikę. Równania dla uzmiennionych stałych można wyprowadzać na dwa zasadnicze sposoby: wychodząc od rozwiązania definiującego (1.17) lub wykorzystując całki układu. To drugie podejście bywa bardziej atrakcyjne, a nie narzuca nam zasadniczych ograniczeń, gdyż metoda uzmienniania stałych wymaga, z założenia, całkowalnego zagadnienia definiującego. Zagadnienie całkowalne to takie, dla którego znamy wszystkie całki ruchu – w naszym przypadku  $2M$  niezależnych całek

$$\Psi_i(\mathbf{r}, \mathbf{v}, t) = S_i = \text{const}, \quad i = 1, \dots, 2M, \quad (1.19)$$

co jest równoważne z możliwością podania ścisłego rozwiązania (1.17). Całkami ruchu nazywa się równie często równania (1.19) jak i same funkcje  $\Psi_i$ .

Całkom ruchu zagadnienia definiującego towarzyszą stałe ruchu  $S_i$ , które albo są wprost stałymi dowolnymi z rozwiązania definiującego, albo znamy związki między  $\mathbf{S}$  i  $\mathbf{C}$ . Ze stałości całki ruchu wynika

$$\dot{S}_i = \dot{\Psi}_i = (D_{\mathbf{r}}\Psi_i) \mathbf{v} + (D_{\mathbf{v}}\Psi_i) \mathbf{F}_0 + D_t\Psi_i = 0, \quad (1.20)$$

gdzie skorzystaliśmy z reguł różniczkowania funkcji złożonej oraz z równań (1.16). Symbolem  $D_{\mathbf{x}}$  oznaczamy transponowany gradient, tzn. jeśli  $\mathbf{x} \in \mathbb{R}^N$  jest wektorem kolumnowym, a  $F(\mathbf{x})$  dowolną funkcją, to

$$D_{\mathbf{x}}F = \left( \frac{\partial F}{\partial x_1}, \dots, \frac{\partial F}{\partial x_N} \right). \quad (1.21)$$

W metodzie uzmienniania stałych, całki zagadnienia definiującego mogą przestać być stałe po wprowadzeniu dodatkowej siły  $\mathbf{P}$ . Zauważmy jednak, że jedynym źródłem zmienności stałych ruchu  $\mathbf{S}$  staje się ta część pochodnej względem czasu, która zawiera  $\mathbf{P}$ , co upoważnia nas do sformułowania równań ruchu

$$\dot{S}_i = (D_{\mathbf{v}}\Psi_i) \mathbf{P}, \quad i = 1, \dots, 2M. \quad (1.22)$$

Wzór ten można stosować „na skróty” według reguły: różniczkujemy względem czasu tylko występujące w całkach  $\mathbf{v}$  i zastępujemy  $\dot{\mathbf{v}}$  siłą  $\mathbf{P}$ .

Jeśli  $\mathbf{S} = \mathbf{C}$ , to równania (1.22) są tymi, których szukaliśmy. W przeciwnym wypadku, musimy je uzupełnić znajomością związków między  $\mathbf{S}$  i  $\mathbf{C}$ , które po zróżniczkowaniu prowadzą do

$$\dot{C}_i = (D_{\mathbf{S}}C_i) \dot{\mathbf{S}}, \quad i = 1, \dots, 2M, \quad (1.23)$$

czyli

$$\dot{\mathbf{C}} = (D_{\mathbf{S}}\mathbf{C}) (D_{\mathbf{v}}\Psi) \mathbf{P}.$$

Jeśli znamy tylko  $\mathbf{S}(\mathbf{C})$ , to będziemy musieli odwrócić macierz Jacobiego:

$$D_{\mathbf{S}}\mathbf{C} = (D_{\mathbf{C}}\mathbf{S})^{-1}.$$

Jest to dość prosty schemat, choć trzeba pamiętać, że prawe strony równań (1.22) muszą zostać wyrażone przy pomocy uzmiennionych stałych  $\mathbf{C}$  za pośrednictwem rozwiązania definiującego (1.17). Jak widać, nawet zastosowanie całek ruchu nie uwalnia nas od wymogu znajomości jawnego rozwiązania.

Rozwiązanie układu (1.22) dostarczy nam  $2M$  funkcji  $C_i(t, \mathbf{C}')$ , które muszą zależeć od nowych stałych dowolnych  $\mathbf{C}'$ . Te zaś najprościej zdefiniować jako warunki początkowe, to znaczy

$$\mathbf{C}' = \mathbf{C}(t_0).$$

Kiedy podstawimy wartości  $\mathbf{C}(t)$  na dany moment czasu do równań (1.17), dostaniemy wartości położenia i prędkości na tę epokę  $\mathbf{r}(t)$  i  $\mathbf{v}(t)$ .

## 1.4 Uzmiennianie stałych dla oscylatora Duffinga

Oscylator Duffinga (1.13) jest układem o jednym stopniu swobody, a więc zamiast wektorów  $\mathbf{r}$  i  $\mathbf{v}$  mamy po prostu  $r = \varphi$  i  $v = \dot{\varphi} = \Phi$ . Układem definiującym z prawą stroną  $F_0 = -\omega_0^2 \varphi$  jest oscylator harmoniczny (1.14), zaś zaburzeniem jest  $P = \frac{1}{6}\omega_0^2 \varphi^3$ . Rozwiązanie definiujące ma więc postać (1.15) zależną od dwóch stałych dowolnych

$$\mathbf{C} = \begin{pmatrix} A \\ \chi \end{pmatrix}.$$

Oscylator harmoniczny posiada dwie całki ruchu. Pierwsza z nich to niezależna od czasu całka energii

$$\Psi_1 = \frac{1}{2} (\Phi^2 + \omega_0^2 \varphi^2) = E = \text{const}, \quad (1.24)$$

ze stałą ruchu  $E$ . Druga całka ruchu musi zależeć od czasu (jeśli układ o  $M$  stopniach swobody ma  $2M$  całek ruchu, to co najmniej jedna z nich musi zależeć od czasu); najprościej wziąć wprost z rozwiązania

$$\Psi_2 = \arctg \left[ \frac{\omega_0 \varphi}{\Phi} \right] - \omega_0 t = \chi = \text{const}, \quad (1.25)$$

ze stałą ruchu  $\chi$ , a więc

$$\mathbf{S} = \begin{pmatrix} E \\ \chi \end{pmatrix}.$$

Zacznijmy od równań (1.22) dla uzmiennionych stałych ruchu po wprowadzeniu siły  $P$ . W przypadku, który rozpatrujemy ( $M = 1$ ), pochodna wektorowa  $D_{\mathbf{v}}$  to zwykła skalarna pochodna  $D_{\Phi}$ , czyli

$$\dot{E} = \frac{\partial \Psi_1}{\partial \Phi} P = \Phi P, \quad (1.26)$$

oraz

$$\dot{\chi} = \frac{\partial \Psi_2}{\partial \Phi} P = \frac{1}{1 + \left(\frac{\omega_0 \varphi}{\Phi}\right)^2} \left(-\frac{\omega_0 \varphi}{\Phi^2}\right) P = -\frac{\omega_0 \varphi}{2E} P. \quad (1.27)$$

Potrzebujemy teraz związków między  $\mathbf{C}$  i  $\mathbf{S}$ , żeby otrzymać równania (1.23). Związki te są proste, gdyż  $C_2 = S_2 = \chi$ , natomiast po wstawieniu (1.15) do całki energii (1.24) widzimy, że

$$E = \frac{\omega_0^2}{2} A^2, \quad (1.28)$$

czyli

$$A \dot{A} = \frac{\dot{E}}{\omega_0^2} = \frac{\Phi P}{\omega_0^2}. \quad (1.29)$$

Ze związku (1.28) możemy skorzystać także w równaniu (1.27) i ostatecznie otrzymujemy równania dla uzmiennionych stałych: najpierw

$$\begin{aligned} \dot{A} &= \frac{\Phi P}{A \omega_0^2}, \\ \dot{\chi} &= -\frac{\varphi P}{A^2 \omega_0}, \end{aligned}$$

a potem, po wprowadzeniu rozwiązania definiującego (1.15),

$$\begin{aligned} \dot{A} &= \frac{P}{\omega_0} \cos(\omega_0 t + \chi), \\ \dot{\chi} &= -\frac{P}{A \omega_0} \sin(\omega_0 t + \chi). \end{aligned} \quad (1.30)$$

Równania (1.30) są ważne dla oscylatora z **dowolną** siłą zaburzającą  $P$ . W przypadku oscylatora Duffinga podstawimy do nich

$$P = \frac{1}{6} \omega_0^2 \varphi^3 = \frac{1}{6} A^3 \omega_0^2 \sin^3(\omega_0 t + \chi). \quad (1.31)$$

Wystarczy już tylko podstawić (1.31) do prawych stron (1.30) a następnie przejść od iloczynów i potęg sinusa i cosinusa do postaci wielomianów trygonometrycznych. Wykorzystamy przy tym

$$(\sin y)^4 = \frac{1}{8} (3 - 4 \cos 2y + \cos 4y),$$

oraz

$$\cos y (\sin y)^3 = \frac{1}{8} (2 \sin 2y - \sin 4y).$$

W ten sposób otrzymujemy końcową postać równań ruchu dla małych drgań wahadła w zmiennych oskulacyjnych

$$\begin{aligned} \dot{A} &= A^3 \omega_0 \left[ \frac{1}{24} \sin 2\psi - \frac{1}{48} \sin 4\psi \right], \\ \dot{\chi} &= A^2 \omega_0 \left[ -\frac{1}{16} + \frac{1}{12} \cos 2\psi - \frac{1}{48} \cos 4\psi \right], \end{aligned} \quad (1.32)$$

gdzie symbol  $\psi$  oznacza

$$\psi = \omega_0 t + \chi.$$

Na pierwszy rzut oka, równania (1.32) wyglądają o wiele bardziej nieliniowo niż pierwotne (1.13). Można się nawet zastanawiać, czy warto je przekształcać do takiej postaci. Odpowiedź na to pytanie zależy od sposobu w jaki będziemy próbowali je rozwiązać. Dopiero gdy zapoznamy się z podstawami metod analitycznych i numerycznych mechaniki nieba stanie się jasne, którą z postaci równań ruchu wygodniej jest zastosować. Spróbujmy jednak pokazać, jak łatwo można teraz znaleźć przybliżone rozwiązanie równań (1.32).

W pierwszym przybliżeniu podstawiamy do prawych stron wartości stałych zagadnienia definiującego  $A_0 = \text{const}$  oraz  $\chi_0 = \text{const}$  wyliczone w epoce  $t_0$  ze zmiennych  $\varphi$  i  $\Phi$ . Mamy wtedy do czynienia z prostymi równaniami

$$\dot{A} \approx f_1(A_0, \chi_0, t), \quad \dot{\chi} \approx f_2(A_0, \chi_0, t),$$

które sprowadzają się do zwykłych całek oznaczonych

$$A \approx A_0 + \int_{t_0}^t f_1 dt, \quad \chi \approx \chi_0 + \int_{t_0}^t f_2 dt.$$

Całkowanie prowadzi do

$$\begin{aligned} A &\approx A_0 \left( 1 - \frac{A_0^2}{48} [\cos 2\psi_0(t) - \cos 2\psi_0(t_0)] \right. \\ &\quad \left. + \frac{A_0^2}{192} [\cos 4\psi_0(t) - \cos 4\psi_0(t_0)] \right), \\ \chi &\approx \chi_0 - \frac{A_0^2 \omega_0}{16} (t - t_0) + \frac{A_0^2}{24} [\sin 2\psi_0(t) - \sin 2\psi_0(t_0)] \\ &\quad - \frac{A_0^2}{192} [\sin 4\psi_0(t) - \sin 4\psi_0(t_0)], \end{aligned}$$

gdzie

$$\psi_0(t) = \omega_0 t + \chi_0.$$

Można sprawdzić, że w epoce oskulacji  $t_0$  mamy faktycznie  $\chi \approx \chi_0$  oraz  $A \approx A_0$ . Jest to tylko pierwsze przybliżenie, ale zawiera cenne informacje: oskulacyjna amplituda drgań oscyluje wokół pewnej średniej wartości, natomiast faza posiada również „wyraz wiekowy” który modyfikuje okres drgań w zależności od amplitudy wychylenia. Uzyskanie podobnych wyników na podstawie oryginalnych równań Duffinga byłoby bardziej złożone.

## 1.5 Elementy oskulacyjne i równania Gaussa

Tam, gdzie ruch orbitalny masy punktowej opisany przez zagadnienie dwóch ciał przestaje być modelem wystarczająco dokładnym, można wprowadzić jako zmienne oskulacyjne uzmiennione stałe orbity keplerowskiej czyli **elementy oskulacyjne orbity**.

Jak wiemy z „Wstępu do mechaniki nieba”, ruch każdej z mas w zagadnieniu dwóch ciał z siłą

$$\mathbf{F}_0 = -\frac{\mu}{r^3} \mathbf{r},$$

jest opisany jednoznacznie przez 6 stałych dowolnych, które najczęściej występują w postaci elementów keplerowskich  $\mathcal{E}$ :

$$a, e, I, \omega, \Omega, M_0 = M(t_0).$$

Znamy już wzory definiujące transformację

$$(\mathbf{r}, \mathbf{v})_t \leftrightarrow \mathcal{E}_{t_0},$$

która z jednej strony pozwala wyliczać na dowolny moment czasu  $t$  położenie i prędkość ciała, dla którego podaliśmy elementy keplerowskie  $\mathcal{E}$  w momencie  $t_0$ , a z drugiej – pozwala wyliczyć z położenia i prędkości w momencie  $t$  elementy  $\mathcal{E}$  dla epoki odniesienia  $t_0$ . Ponieważ  $\mathcal{E}$  są stałymi ruchu, to biorąc  $(\mathbf{r}, \mathbf{v})$  w dowolnym momencie  $t$  i wyliczając z nich elementy otrzymamy zawsze te same wartości  $\mathcal{E}$ .

Jeśli pojawi się dodatkowa siła i ruch ciała nie jest już ruchem keplerowskim, to nadal możemy wyliczać elementy keplerowskie według znanych już wzorów, ale teraz w każdej epoce  $t$  możemy otrzymać inne wartości elementów  $\mathcal{E}_{t_0}$  dla epoki odniesienia  $t_0$ . Zgodnie z zasadą metody uzmienniania stałych zaczynamy traktować elementy orbity jako funkcje czasu i nazywamy je **elementami oskulacyjnymi**.

Należy wyraźnie odróżnić dwa momenty czasu pojawiające się w tym opisie. **Epoka oskulacji**  $t$  jest momentem czasu, w którym mamy dane położenie i prędkość służące do wyliczenia elementów  $\mathcal{E}$ . Jednym z tych elementów jest **anomalii średnia epoki odniesienia**  $M_0 = M(t_0)$ , którą otrzymujemy „cofając” lub „popychając” wartość anomalii średniej  $M$  z epoki  $t$  do epoki  $t_0$ . Ponieważ uzmienniliśmy stałe, to  $M_0$  może być zmienną i w ogólności dla różnych epok oskulacji  $t_1$  i  $t_2$  będziemy mieli  $M_0(t_1) \neq M_0(t_2)$ .

Używamy więc pojęcia  $M_0(t)$  – anomalii średniej epoki odniesienia  $t_0$  dla epoki oskulacji  $t$ . Czasami można uprościć sobie życie przyjmując, że  $t = t_0$ . W przypadku pozostałych elementów nie mamy tego problemu i wystarcza używanie pojęcia epoki oskulacji dla  $a(t), e(t)$ , itd.

Mówiąc najkrócej, elementy oskulacyjne epoki  $t$  to wartości elementów keplerowskich wyliczone dla dowolnej trajektorii wzorami pochodzącymi z zagadnienia dwóch ciał. Jeśli w poprzednim rozdziale traktowaliśmy wahadło jak oscylator o zmiennej w czasie amplitudzie i fazie, to teraz będziemy traktować dowolną trajektorię jak orbitę keplerowską o zmiennych w czasie elementach.

Inna, bardziej pogładowa definicja brzmi: elementy oskulacyjne epoki  $t$ , to elementy orbity, po której poruszałoby się ciało, gdyby w epoce  $t$  przestała działać siła zaburzająca.

Przystąpimy teraz do sformułowania równań ewolucji elementów oskulacyjnych dla dowolnej siły  $\mathbf{P}$  zaburzającej zagadnienie dwóch ciał

$$\begin{aligned}\dot{\mathbf{r}} &= \mathbf{v}, \\ \dot{\mathbf{v}} &= -\frac{\mu}{r^3}\mathbf{r} + \mathbf{P}.\end{aligned}\tag{1.33}$$

Znaczenie stałej  $\mu$  zależy od typu zagadnienia – względnego lub barycentrycznego. Wyprowadzenie równań dla  $\dot{a}$ ,  $\dot{e}$ , itd. przy dowolnej postaci siły  $\mathbf{P}$  – zwanych równaniami Gaussa – przeprowadzimy wykorzystując znane całki ruchu zagadnienia dwóch ciał.

### 1.5.1 Uzmiennione całki ruchu

Przypomnijmy całki ruchu pojawiające się we względnym lub barycentrycznym zagadnieniu dwóch ciał. Bez względu na typ ruchu, mamy 7 całek niezależnych od czasu:

- Całkę siły żywej *vel* energii

$$h = \frac{1}{2}\mathbf{v} \cdot \mathbf{v} - \frac{\mu}{\sqrt{\mathbf{r} \cdot \mathbf{r}}}.\tag{1.34}$$

- Całki pól (momentu pędu)

$$\mathbf{G} = \mathbf{r} \times \mathbf{v}.\tag{1.35}$$

- Całki Laplace'a

$$\mathbf{e} = \frac{\mathbf{v} \times \mathbf{G}}{\mu} - \frac{\mathbf{r}}{r}, \quad (1.36)$$

które można także przedstawić w postaci

$$\mu \mathbf{e} = \left( v^2 - \frac{\mu}{r} \right) \mathbf{r} - (\mathbf{r} \cdot \mathbf{v}) \mathbf{v}. \quad (1.37)$$

Tylko 5 spośród nich jest wzajemnie niezależnych, gdyż istnieją dwa związki:  $\mathbf{G} \cdot \mathbf{e} = 0$ , oraz  $e^2 = 1 + 2hG^2\mu^{-2}$ . Szóstą całką ruchu jest, w zależności od typu orbity, równanie Keplera (eliptyczne lub hiperboliczne) albo równanie Barkera. Ograniczymy nasze rozważania do orbit eliptycznych jako rozwiązania definiującego, więc szóstą całką ruchu będzie

$$M_0 = E - e \sin E - n(t - t_0), \quad (1.38)$$

gdzie ruch średni  $n$  (nadal definiowany poprzez III prawo Keplera  $n^2 a^3 = \mu$ ) i anomalia mimośrodowa  $E$  są znanymi funkcjami położenia i prędkości.

Zastosujmy teraz do pierwszych siedmiu całek wzór (1.22), a otrzymamy

$$\dot{h} = \mathbf{v} \cdot \mathbf{P}, \quad (1.39)$$

$$\dot{\mathbf{G}} = \mathbf{r} \times \mathbf{P}, \quad (1.40)$$

$$\begin{aligned} \mu \dot{\mathbf{e}} &= 2(\mathbf{v} \cdot \mathbf{P})\mathbf{r} - (\mathbf{r} \cdot \mathbf{P})\mathbf{v} - (\mathbf{r} \cdot \mathbf{v})\mathbf{P} = \\ &= \mathbf{P} \times \mathbf{G} + \mathbf{v} \times (\mathbf{r} \times \mathbf{P}). \end{aligned} \quad (1.41)$$

Ze względu na wyższy poziom komplikacji, równanie Keplera zostawimy sobie „na później”.

### 1.5.2 Równanie dla pólki wielkiej

Równania (1.39-1.41) opisują zmiany stałych ruchu, które nie są elementami Keplerskimi. Są przy tym na tyle ogólne, że nawet trochę żal je psuć przejściem do elementów oskulacyjnych. Na przykład równanie dla  $\dot{h}$  jest ważne dla wszystkich typów orbit, ale wydobycie z niego równania na zmiany oskulacyjnej pólki wielkiej  $a$  ograniczy jego stosowanie do orbit eliptycznych. Mamy bowiem wzór

$$h = -\frac{\mu}{2a}, \quad (1.42)$$

który wiąże stałą siły żywej  $h$  z pólką wielką elipsy, ale dla orbit hiperbolicznych, gdzie mamy pólkę rzeczywistą  $a$ , prawa strona ma inny znak, natomiast dla orbit parabolicznych mamy z definicji  $h = 0$  i dopuszczenie zmian  $h$  jest samo w sobie problematyczne.



Ograniczmy się zatem do zaburzonego ruchu eliptycznego i wtedy

$$\dot{h} = \frac{\mu}{2a^2} \dot{a}. \quad (1.43)$$

Przyrównując stronami (1.39) oraz (1.43) otrzymujemy pierwsze z równań Gaussa

$$\dot{a} = \frac{2a^2}{\mu} \mathbf{v} \cdot \mathbf{P}. \quad (1.44)$$

### 1.5.3 Równanie dla mimośrod

Mimośród orbity to długość wektora Laplace'a  $\mathbf{e}$ . A zatem równanie opisujące zmiany oskulacyjnego mimośrod możemy uzyskać z równania (1.41) dla  $\dot{\mathbf{e}}$ . Wystarczy w tym celu skorzystać z pożytecznej tożsamości: dla każdego wektora  $\mathbf{x}$  obowiązuje

$$\mathbf{x} \cdot \dot{\mathbf{x}} = x\dot{x}.$$

A jeśli użyjemy wektora  $\hat{\mathbf{x}} = \mathbf{x}/|\mathbf{x}| = \mathbf{x}/x$ , o jednostkowej długości, to

$$\hat{\mathbf{x}} \cdot \dot{\mathbf{x}} = \dot{x}.$$

Pomnożmy więc obie strony (1.41) skalarnie przez wektor  $\hat{\mathbf{e}}$ . Po lewej stronie pojawi się  $\hat{\mathbf{e}} \cdot \dot{\mathbf{e}} = \dot{e}$  i otrzymamy drugie z równań Gaussa

$$\mu \dot{e} = 2(\mathbf{r} \cdot \hat{\mathbf{e}})(\mathbf{v} \cdot \mathbf{P}) - (\mathbf{v} \cdot \hat{\mathbf{e}})(\mathbf{r} \cdot \mathbf{P}) - (\mathbf{r} \cdot \mathbf{v})(\hat{\mathbf{e}} \cdot \mathbf{P}). \quad (1.45)$$

Ograniczenie w stosowaniu tego równania pojawiło się w sposób nieco zakamufłowany: traci ono sens dla orbit kołowych z  $\mathbf{e} = \mathbf{0}$ , gdyż wtedy nie istnieje wektor  $\hat{\mathbf{e}}$ .

### 1.5.4 Kąty Eulera a prędkość kątowna

Po znalezieniu wzorów dla  $\dot{a}$  i  $\dot{e}$  stajemy wobec problemu powiązania zmian wektorów  $\mathbf{G}$  i  $\mathbf{e}$  ze zmianami kolejnych elementów keplerowskich:  $\Omega$ ,  $I$ ,  $\omega$ , które mają charakter kątów Eulera typu 3-1-3 orientujących orbitę w przestrzeni. Wykorzystamy do tego ważne twierdzenie dotyczące związku prędkości kątowej z obrotami o kąty Eulera typu 3-1-3.

Rozpatrzmy dwa prawoskrętne układy współrzędnych: stały, o wektorach osi  $(\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}})$  i obracający się, o wektorach osi  $(\hat{\mathbf{i}}'', \hat{\mathbf{j}}'', \hat{\mathbf{k}}'')$ . W każdym momencie czasu możemy powiązać te dwa układy przez chwilowe wartości kątów Eulera  $\phi, \vartheta, \psi$ , związane z trzema obrotami:

- Wokół osi  $\hat{\mathbf{k}}$  podstawowego układu współrzędnych  $(\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}})$ , o kąt  $\phi$ . Taki obrót przenosi wektor  $\hat{\mathbf{i}}$  w położenie  $\hat{\mathbf{i}}'$  wyznaczające linię przecięcia płaszczyzn podstawowych obu układów.
- Obrót wokół chwilowej osi  $\hat{\mathbf{i}}'$  o kąt  $\vartheta$ . W wyniku tej operacji wektor  $\hat{\mathbf{k}}$ , który nie uległ zmianie przy pierwszym obrocie, przechodzi w nowy wektor  $\hat{\mathbf{k}}''$ .
- Obrót wokół osi  $\hat{\mathbf{k}}''$  o kąt  $\psi$ , który przenosi wektor  $\hat{\mathbf{i}}'$  w wektor  $\hat{\mathbf{i}}''$ .

Jeśli kąty Eulera zmieniają się w czasie w sposób ciągły, to punkt wyznaczony przez koniec radialnego wektora  $\hat{\mathbf{r}}$  sztywno związanego z układem  $(\hat{\mathbf{i}}'', \hat{\mathbf{j}}'', \hat{\mathbf{k}}'')$  (takiego, którego rzuty  $\hat{\mathbf{r}} \cdot \hat{\mathbf{i}}''$ ,  $\hat{\mathbf{r}} \cdot \hat{\mathbf{j}}''$ ,  $\hat{\mathbf{r}} \cdot \hat{\mathbf{k}}''$  są stałe – w szczególności jeden z wektorów osi) zyskuje prędkość kątową  $\boldsymbol{\omega}_r$  i mamy względem osi stałych

$$\frac{d\hat{\mathbf{r}}}{dt} = \boldsymbol{\omega}_r \times \hat{\mathbf{r}}. \quad (1.46)$$

**TWIERDZENIE 1** Prędkość kątowa obrotu opisanego kątami Eulera 3-1-3  $\phi, \vartheta, \psi$  wynosi

$$\boldsymbol{\omega}_r = \dot{\phi} \hat{\mathbf{k}} + \dot{\vartheta} \hat{\mathbf{i}}' + \dot{\psi} \hat{\mathbf{k}}''. \quad (1.47)$$

Dowód tego twierdzenia znaleźć można w podręcznikach mechaniki klasycznej, ale – jak zauważył C. Leubner (1981, Am. J. Phys. 49, 323) – „proste” dowody wzoru (1.47), które trafiły do podręczników, są zazwyczaj niepoprawne.

Warto zapamiętać, że podane twierdzenie jest ważne nie tylko w mechanice. W astronomii sferycznej i astrometrii może ono być zastosowane do opisu wpływu małych zmian kątów Eulera na położenie punktu na sferze niebieskiej

$$\Delta \hat{\mathbf{r}} \approx \left( \Delta \phi \hat{\mathbf{k}} + \Delta \vartheta \hat{\mathbf{i}}' + \Delta \psi \hat{\mathbf{k}}'' \right) \times \hat{\mathbf{r}}.$$

## WYKŁAD 4

### 1.5.5 Równania dla oskulacyjnej długości węzła wstępującego i nachylenia

Przypomnijmy, że wektor momentu pędu  $\mathbf{G}$  definiuje orientację płaszczyzny orbity w przestrzeni. Jest on prostopadły do tej płaszczyzny, więc nachylenie  $I$  definiujemy poprzez iloczyn skalarny z wersorem osi  $z$  układu

$$\hat{\mathbf{G}} \cdot \hat{\mathbf{z}} = \cos I, \quad (1.48)$$

natomiast kierunek do węzła wstępującego dany jest wersorem

$$\hat{\mathbf{m}} = \frac{\hat{\mathbf{z}} \times \hat{\mathbf{G}}}{\sin I}. \quad (1.49)$$

Ponieważ kąt między wersorem osi  $x$  a wersorem  $\hat{\mathbf{m}}$  jest równy długości węzła wstępującego  $\Omega$ , widzimy że równania dla zmian  $I$  oraz  $\Omega$  powinny wynikać z równań (1.40) dla  $\dot{\mathbf{G}}$ .

Przypomnijmy teraz definicję układu orbitalnego perycentrycznego, którego trzy wersory to:  $\hat{\mathbf{e}}$ ,  $\hat{\mathbf{Q}} = (\hat{\mathbf{G}} \times \hat{\mathbf{e}})$ , oraz  $\hat{\mathbf{G}}$ . Jest on powiązany z układem dowolnym o wersorach  $\hat{\mathbf{x}}$ ,  $\hat{\mathbf{y}}$ ,  $\hat{\mathbf{z}}$  obrotami o kąty  $\phi = \Omega$ ,  $\vartheta = I$ , oraz  $\psi = \omega$ , a kolejne osie obrotu to  $\hat{\mathbf{k}} = \hat{\mathbf{z}}$ ,  $\hat{\mathbf{i}}' = \hat{\mathbf{m}}$ ,  $\hat{\mathbf{k}}'' = \hat{\mathbf{G}}$ .

Przedstawmy moment pędu  $\mathbf{G}$  jako iloczyn długości i wersora  $G\hat{\mathbf{G}}$ . Równanie (1.40) przybiera wtedy postać

$$\dot{\mathbf{G}} = \frac{d}{dt} (G \hat{\mathbf{G}}) = \dot{G} \hat{\mathbf{G}} + G \frac{d\hat{\mathbf{G}}}{dt} = \mathbf{r} \times \mathbf{P}.$$

Skąd biorą się zmiany kierunku  $\hat{\mathbf{G}}$ ? Możemy uznać, że wersor  $\mathbf{G}$ , jest to wektor  $\hat{\mathbf{z}}$  aktywnie obrócony o kąty Eulera  $\phi = \Omega$  i  $\vartheta = I$  (trzeci kąt  $\psi = \omega$  jest nieistotny). Ciągła zmiany kątów Eulera  $\Omega$ ,  $I$  wywołują prędkość katową końca  $\hat{\mathbf{G}}$  i stosując Twierdzenie 1 uzyskamy wektorową równość

$$\dot{G} \hat{\mathbf{G}} + G \sin I \dot{\Omega} \hat{\mathbf{m}} + G \dot{I} \hat{\mathbf{m}} \times \hat{\mathbf{G}} = \mathbf{r} \times \mathbf{P}, \quad (1.50)$$

która nie posiada wyrazu związanego z  $\dot{\omega}$ , gdyż  $G \dot{\omega} \hat{\mathbf{G}} \times \hat{\mathbf{G}} = \mathbf{0}$ . W wyrazie z  $\dot{\Omega}$  wykorzystaliśmy fakt, że  $\hat{\mathbf{z}} \times \hat{\mathbf{G}} = \sin I \hat{\mathbf{m}}$ , w myśl równania (1.49).

Jak z układu trzech równań (1.50) otrzymać dwa osobne równania skalarnie dla  $\dot{\Omega}$  oraz  $\dot{I}$ ? Najgorsze, co by można wymyślić, to rozpisać wszystkie wektory na składowe w wybranej bazie i rozwiązywać układ trzech równań z

niewiadomymi  $\dot{G}$ ,  $\dot{\Omega}$ ,  $\dot{I}$ . O wiele bardziej wydajne jest jednak podejście wektorowe, polegające na wykonaniu iloczynu skalarnego z umiejętnie dobranym wektorem.

Jeśli interesuje nas  $\dot{\Omega}$ , to powinniśmy pomnożyć obie strony (1.50) przez wektor który jest prostopadły zarówno do  $\hat{\mathbf{G}}$ , jak i do  $\hat{\mathbf{m}} \times \hat{\mathbf{G}}$ . Z tego punktu widzenia, idealnym wyborem jest wersor  $\hat{\mathbf{m}}$ , który prowadzi do

$$\dot{G} \hat{\mathbf{m}} \cdot \hat{\mathbf{G}} + G \sin I \dot{\Omega} \hat{\mathbf{m}} \cdot \hat{\mathbf{m}} + G \dot{I} \hat{\mathbf{m}} \cdot (\hat{\mathbf{m}} \times \hat{\mathbf{G}}) = \hat{\mathbf{m}} \cdot (\mathbf{r} \times \mathbf{P}).$$

Ponieważ  $\hat{\mathbf{m}} \cdot \hat{\mathbf{G}} = 0$ ,  $\hat{\mathbf{m}} \cdot \hat{\mathbf{m}} = 1$ , oraz w myśl reguły cyklicznego przestawiania czynników w mieszanym iloczynie wektorowym,

$$\hat{\mathbf{m}} \cdot (\hat{\mathbf{m}} \times \hat{\mathbf{G}}) = \hat{\mathbf{G}} \cdot (\hat{\mathbf{m}} \times \hat{\mathbf{m}}) = 0,$$

zostaje nam

$$G \sin I \dot{\Omega} = \mathbf{P} \cdot (\hat{\mathbf{m}} \times \mathbf{r}),$$

gdzie prawa strona powstała przez cykliczne przestawienie czynników. Ponieważ  $\hat{\mathbf{m}}$  i  $\mathbf{r}$  leżą w płaszczyźnie orbity i ich iloczyn wektorowy ma kierunek  $\hat{\mathbf{G}}$ , a kąt jaki tworzą to tzw. argument szerokości  $f + \omega$  (suma anomalii prawdziwej  $f$  i argumentu perycentrum), dochodzimy do kolejnego z równań Gaussa

$$G \sin I \dot{\Omega} = \mathbf{P} \cdot \hat{\mathbf{G}} r \sin(f + \omega). \quad (1.51)$$

W podobny sposób otrzymamy równanie dla zmian nachylenia. Tym razem optymalny wybór wektora do iloczynu skalarnego to wersor  $(\hat{\mathbf{m}} \times \hat{\mathbf{G}})$ , prostopadły zarówno do linii węzłów, jak i do momentu pędu. Mnożąc (1.50) skalarnie przez ten wektor, otrzymujemy

$$G \dot{I} = (\hat{\mathbf{m}} \times \hat{\mathbf{G}}) \cdot (\mathbf{r} \times \mathbf{P}) = \mathbf{P} \cdot ((\hat{\mathbf{m}} \times \hat{\mathbf{G}}) \times \mathbf{r}).$$

Zależnie od stopnia posiadanej wyobraźni przestrzennej, albo stosujemy tożsamość „bac-cab”

$$(\hat{\mathbf{m}} \times \hat{\mathbf{G}}) \times \mathbf{r} = (\hat{\mathbf{m}} \cdot \mathbf{r}) \hat{\mathbf{G}} - (\hat{\mathbf{G}} \cdot \mathbf{r}) \hat{\mathbf{m}} = r \cos(f + \omega) \hat{\mathbf{G}},$$

albo analizujemy kierunki i kąty między poszczególnymi wektorami, otrzymując ten sam wynik

$$G \dot{I} = \mathbf{P} \cdot \hat{\mathbf{G}} r \cos(f + \omega). \quad (1.52)$$

Obydwa równania (1.51) oraz (1.52) nie mają sensu dla  $I = 0$  lub  $I = \pi$ , gdyż korzystaliśmy w wyprowadzeniu z wersora  $\hat{\mathbf{m}}$ , który nie istnieje w tych przypadkach. Również dla orbit prostoliniowych, gdy  $\mathbf{G} = \mathbf{0}$ , równania te tracą sens, gdyż nie istniałby wersor  $\hat{\mathbf{G}}$ .

### 1.5.6 Równanie dla oskulacyjnego argumentu perycentrum

Zmiany argumentu perycentrum, czyli kąta między  $\hat{\mathbf{m}}$  a  $\hat{\mathbf{e}}$ , otrzymamy z uzmiennionych całek Laplace'a (1.41). Podobnie jak dla momentu pędu, przedstawimy wektor Laplace'a jako iloczyn  $\mathbf{e} = e\hat{\mathbf{e}}$  i uznamy  $\hat{\mathbf{e}}$  za wersor  $\hat{\mathbf{x}}$  obrócony o kąty  $\Omega, I, \omega$  po czym skorzystamy z Twierdzenia 1

$$\dot{\mathbf{e}} = \dot{e}\hat{\mathbf{e}} + e\frac{d\hat{\mathbf{e}}}{dt} = \dot{e}\hat{\mathbf{e}} + e\left(\dot{\Omega}\hat{\mathbf{z}} + \dot{I}\hat{\mathbf{m}} + \dot{\omega}\hat{\mathbf{G}}\right) \times \hat{\mathbf{e}}. \quad (1.53)$$

Niestety, nie możemy liczyć na usunięcie wszystkich pochodnych oprócz  $\dot{\omega}$ , więc stosujemy wariant skromniejszy: iloczyn skalarny z wersorem

$$\hat{\mathbf{Q}} = \hat{\mathbf{G}} \times \hat{\mathbf{e}},$$

który leży w płaszczyźnie orbity i jest odległy od perycentrum o  $90^\circ$  w kierunku ruchu orbitalnego. Z definicji,  $\hat{\mathbf{Q}} \cdot \hat{\mathbf{e}} = 0$ , więc podstawiając (1.53) do (1.41) i mnożąc obie strony przez  $\hat{\mathbf{Q}}$  mamy

$$\mu e \left( \dot{\Omega} \hat{\mathbf{Q}} \cdot (\hat{\mathbf{z}} \times \hat{\mathbf{e}}) + \dot{I} \hat{\mathbf{Q}} \cdot (\hat{\mathbf{m}} \times \hat{\mathbf{e}}) + \dot{\omega} \right) = \hat{\mathbf{Q}} \cdot \mathbf{R},$$

gdzie przez  $\mathbf{R}$  oznaczyliśmy jedną z dwóch postaci prawych stron równania wektorowego (1.41). Łatwo sprawdzić, że

$$\hat{\mathbf{Q}} \cdot (\hat{\mathbf{z}} \times \hat{\mathbf{e}}) = \hat{\mathbf{z}} \cdot (\hat{\mathbf{e}} \times \hat{\mathbf{Q}}) = \hat{\mathbf{z}} \cdot \hat{\mathbf{G}} = \cos I,$$

oraz

$$\hat{\mathbf{Q}} \cdot (\hat{\mathbf{m}} \times \hat{\mathbf{e}}) = \hat{\mathbf{m}} \cdot (\hat{\mathbf{e}} \times \hat{\mathbf{Q}}) = \hat{\mathbf{m}} \cdot \hat{\mathbf{G}} = 0,$$

więc, wybierając pierwszą postać prawych stron,

$$\mu e \left( \dot{\omega} + \cos I \dot{\Omega} \right) = 2(\mathbf{r} \cdot \hat{\mathbf{Q}})(\mathbf{v} \cdot \mathbf{P}) - (\mathbf{v} \cdot \hat{\mathbf{Q}})(\mathbf{r} \cdot \mathbf{P}) - (\mathbf{r} \cdot \mathbf{v})(\hat{\mathbf{Q}} \cdot \mathbf{P}). \quad (1.54)$$

Wyraz zawierający  $\dot{\Omega}$  zostawiliśmy chwilowo po lewej stronie dla podkreślenia, że przy  $I = 0$  (gdy  $\cos I = 1$ ) prowadzi on do dobrze określonego równania dla długości perycentrum  $\dot{\varpi} = \dot{\omega} + \dot{\Omega}$ . Natomiast w przypadku  $I = \pi$ ,  $e = 0$  lub orbit prostoliniowych, równanie Gaussa (1.54) traci sens.

## WYKŁAD 5

### 1.5.7 Równania dla zmian anomalii

Pozostało nam już tylko równanie dla zmian oskulacyjnej anomalii średniej epoki odniesienia  $t_0$ , czyli  $\dot{M}_0$  w zagadnieniu zaburzonym. Trafiamy tutaj na dość nieprzyjemny problem. W pierwotnej postaci, równanie Keplera (1.38) miało postać

$$M = E - e \sin E.$$

Anomalia średnia jest w ruchu keplerowskim dana poprzez

$$M(t) = n [t - t_0] + M_0.$$

Tylko że w zagadnieniu dwóch ciał ruch średni  $n = \sqrt{\mu a^{-3}}$  jest stały, podczas gdy w ruchu zaburzonym zmienność pólوسی wielkiej uzmiennia  $n$ . Jak więc traktować  $M$  ? Czy przyjąć

$$M(t) = n(t) [t - t_0] + M_0, \quad (1.55)$$

czy może

$$M(t) = \int_{t_0}^t n(t) dt + M_0, \quad (1.56)$$

jako definicję podstawową ? Każde z tych podejść prowadzi do innej definicji  $M_0$  w przypadku zaburzonym. Wariant (1.56) przyjmowany jest najczęściej, gdyż prowadzi do względnie prostego

$$\dot{M} = n(t) + \dot{M}_0, \quad (1.57)$$

podczas gdy (1.55) oznacza

$$\dot{M} = n + \dot{n} [t - t_0] + \dot{M}_0, \quad (1.58)$$

co prowadzi do nieprzyjemnej sytuacji: jeśli  $\dot{n}$  ma charakter ograniczonych oscylacji o częstotliwości  $\alpha$ , to pojawią się bardzo niepożądane wyrazy typu  $t \cos \alpha t$ , zwane mieszanymi wyrazami wiekowymi.

Żeby uniknąć dylematów związanych z definicją  $M_0$ , obejdźmy problem w ten sposób, że wyprowadzimy równanie dla  $\dot{M}$ , a wybór definicji  $M_0$  pozostawimy otwarty. Co więcej, w wielu sytuacjach można się bez niego obejść i przestać na  $M(t)$ .

Różniczkując równanie Keplera w przypadku zaburzonym

$$\dot{M} = \dot{E}(1 - e \cos E) - \dot{e} \sin E, \quad (1.59)$$

widzimy, że potrzebna będzie pochodna anomalii mimośrodowej, gdyż wzór na  $\dot{e}$  został wyprowadzony wcześniej. Informację o  $\dot{E}$  najlepiej otrzymać na podstawie prędkości kątowej anomalii prawdziwej  $\dot{f}$ , gdyż ta druga ma prostą definicję geometryczną jako kąt między  $\hat{e}$  i  $\hat{r}$ .

Wyjdźmy od oczywistego wzoru na prędkość

$$\mathbf{v} = \dot{\mathbf{r}} = \frac{d(r \hat{\mathbf{r}})}{dt} = \dot{r} \hat{\mathbf{r}} + r \frac{d\hat{\mathbf{r}}}{dt}.$$

Zuważmy teraz, że wektor położenia  $\hat{\mathbf{r}}$  możemy otrzymać z wektora  $\hat{\mathbf{x}}$  poprzez obrót o kąty Eulera  $\Omega$ ,  $I$ ,  $\omega + f$  ze znajomymi wektorami obrotu  $\hat{\mathbf{z}}$ ,  $\hat{\mathbf{m}}$ ,  $\hat{\mathbf{G}}$ . A więc można zastosować Twierdzenie 1 i napisać

$$\dot{r} \hat{\mathbf{r}} + r \left( \dot{\Omega} \hat{\mathbf{z}} + \dot{I} \hat{\mathbf{m}} + (\dot{\omega} + \dot{f}) \hat{\mathbf{G}} \right) \times \hat{\mathbf{r}} = \mathbf{v}.$$

Pomnóżmy skalarnie obie strony przez wektor

$$\hat{\mathbf{t}} = (\hat{\mathbf{G}} \times \hat{\mathbf{r}}),$$

prostopadły do wektora położenia, leżący w płaszczyźnie orbity i skierowany zgodnie z ruchem ciała. Po uwzględnieniu  $\hat{\mathbf{r}} \cdot \hat{\mathbf{t}} = 0$ , oraz

$$\hat{\mathbf{t}} \cdot (\hat{\mathbf{m}} \times \hat{\mathbf{r}}) = \hat{\mathbf{m}} \cdot (\hat{\mathbf{r}} \times \hat{\mathbf{t}}) = \hat{\mathbf{m}} \cdot \hat{\mathbf{G}} = 0,$$

$$\hat{\mathbf{t}} \cdot (\hat{\mathbf{z}} \times \hat{\mathbf{r}}) = \hat{\mathbf{z}} \cdot (\hat{\mathbf{r}} \times \hat{\mathbf{t}}) = \hat{\mathbf{z}} \cdot \hat{\mathbf{G}} = \cos I,$$

otrzymujemy

$$r \left( \dot{\Omega} \cos I + \dot{\omega} + \dot{f} \right) = \mathbf{v} \cdot \hat{\mathbf{t}}.$$

Ale to, co widzimy po prawej stronie, to nic innego jak prędkość transwersalna, która z defincji ma postać

$$v_t = \mathbf{v} \cdot \hat{\mathbf{t}} = \frac{G}{r}.$$

Wzór dla prędkości kątowej anomalii prawdziwej w ruchu zaburzonym ma zatem postać

$$\dot{f} = \frac{G}{r^2} - \dot{\omega} - \cos I \dot{\Omega}, \quad (1.60)$$

taką jak w ruchu keplerowskim, tylko z poprawką na ruch linii apsyd od której mierzymy  $f$ .

Aby znaleźć związek między  $\dot{E}$  a  $\dot{f}$  w przypadku zaburzonym, zróżniczkujemy funkcję

$$\frac{r}{a} = 1 - e \cos E = \frac{1 - e^2}{1 + e \cos f},$$

która może służyć jako związek między  $f$  i  $E$ . Uwzględniając możliwość zmian mimośrod, dostaniemy

$$-\cos E \dot{e} + e \sin E \dot{E} = -\frac{r}{p} \left[ 2e + \frac{r}{a} \cos f \right] \dot{e} + \frac{e r^2 \sin f}{a p} \dot{f}, \quad (1.61)$$

gdzie  $p = a(1 - e^2)$ , dla orbit eliptycznych. Po niezbyt wyrafinowanych przekształceniach, uwzględniających

$$r = a(1 - e \cos E), \quad r \sin f = a\sqrt{1 - e^2} \sin E, \quad r \cos f = a(\cos E - e),$$

równanie (1.61) przyjmuje postać

$$\dot{E} = \frac{r}{a\sqrt{1 - e^2}} \dot{f} - \frac{\sin E}{1 - e^2} \dot{e}. \quad (1.62)$$

Podstawiając (1.62) i (1.60) do (1.59) dostajemy

$$\dot{M} = n - \frac{r^2}{a^2\sqrt{1 - e^2}} (\dot{\omega} + \cos I \dot{\Omega}) - \left( 1 + \frac{r}{p} \right) \sin E \dot{e}. \quad (1.63)$$

Uznamy to równanie za szóste z równań Gaussa, gdyż stosownie do wyboru definicji  $M_0$  prowadzi ono bezpośrednio do  $\dot{M}_0$  w połączeniu z równaniem (1.58) lub (1.57).

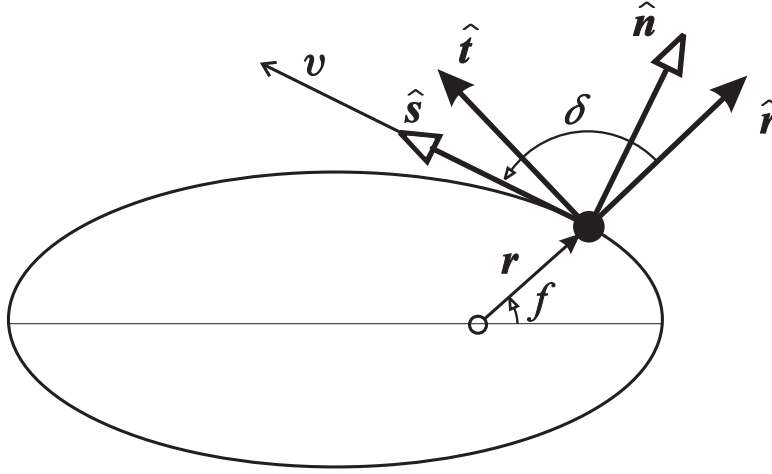
### 1.5.8 Składowe siły zaburzającej

Przyglądając się równaniom (1.44, 1.45, 1.51, 1.52, 1.54) stwierdzamy, że zawierają one iloczyny skalarne siły zaburzającej  $\mathbf{P}$  z różnymi wektorami:  $\mathbf{r}$ ,  $\mathbf{v}$ ,  $\hat{\mathbf{e}}$ ,  $\mathbf{Q}$  i  $\hat{\mathbf{G}}$ . Pierwsze cztery z nich leżą w płaszczyźnie oskulacyjnej orbity, prostopadłej do chwilowego momentu pędu  $\mathbf{G}$ . Staramy się ograniczyć liczbę tych iloczynów rozkładając  $\mathbf{P}$  na składowe w jednej z trzech podstawowych baz: radialnej, stycznej lub perycentrycznej. Wszystkie trzy bazy mają wspólny wersor  $\hat{\mathbf{G}}$ , który dalej oznaczamy będziemy przez  $\hat{\mathbf{b}}$  i nazywać, zgodnie z terminologią geometrii różniczkowej, wersorem binormalnym.

Baza radialna zdefiniowana jest przez wersory

- radialny  $\hat{\mathbf{r}}$  – wzdłuż promienia wodzącego,
- transwersalny  $\hat{\mathbf{t}}$  – prostopadły do radialnego, leżący w płaszczyźnie orbity i skierowany zgodnie z ruchem ciała,
- binormalny (normalny do płaszczyzny orbity)  $\hat{\mathbf{b}}$  – wzdłuż wektora momentu pędu.





Rysunek 1.2: Bazy radialna i styczna. Wersor binormalny skierowany jest „do góry” – prostopadle do płaszczyzny rysunku.

Składowe siły  $\mathbf{P}$  w tej bazie oznaczymy odpowiednio przez  $R$ ,  $T$  i  $B$ :

$$\mathbf{P} = R\hat{\mathbf{r}} + T\hat{\mathbf{t}} + B\hat{\mathbf{b}}.$$

Ich wartości można zdefiniować przy pomocy iloczynów  $\mathbf{P}$  z wektorami  $\mathbf{r}$  i  $\mathbf{G} = \mathbf{r} \times \mathbf{v}$

$$\begin{aligned} R &= \mathbf{P} \cdot \hat{\mathbf{r}} = \mathbf{P} \cdot \frac{\mathbf{r}}{r}, \\ T &= \mathbf{P} \cdot \hat{\mathbf{t}} = \mathbf{P} \cdot \left( \frac{\mathbf{G}}{G} \times \frac{\mathbf{r}}{r} \right), \\ B &= \mathbf{P} \cdot \hat{\mathbf{b}} = \mathbf{P} \cdot \frac{\mathbf{G}}{G}. \end{aligned} \quad (1.64)$$

Baza styczna składa się z następujących wersorów:

- normalnego  $\hat{\mathbf{n}}$ , który określa kierunek normalnej zewnętrznej do krzywej jaką jest orbita,
- stycznego  $\hat{\mathbf{s}}$ , który pokrywa się z kierunkiem wektora prędkości,
- binormalnego  $\hat{\mathbf{b}}$ , wspólnego dla obu baz.

Siła  $\mathbf{P}$  ma w tej bazie składowe  $N$ ,  $S$  i  $B$ :

$$\mathbf{P} = N\hat{\mathbf{n}} + S\hat{\mathbf{s}} + B\hat{\mathbf{b}}.$$

Wartości tych składowych można wyliczyć z definicji

$$\begin{aligned} N &= \mathbf{P} \cdot \hat{\mathbf{n}} = \mathbf{P} \cdot \left( \frac{\mathbf{v}}{v} \times \frac{\mathbf{G}}{G} \right), \\ S &= \mathbf{P} \cdot \hat{\mathbf{s}} = \mathbf{P} \cdot \frac{\mathbf{v}}{v}, \\ B &= \mathbf{P} \cdot \hat{\mathbf{b}} = \mathbf{P} \cdot \frac{\mathbf{G}}{G}. \end{aligned} \quad (1.65)$$

Ponieważ składowa  $B$  jest jednakowa w obu bazach, pozostaje nam jedynie problem związku między składowymi  $R$ ,  $T$  i  $N$ ,  $S$ .

Transformacja z r-t-b do n-s-b jest elementarna, gdyż wymaga jedynie obrotu o kąt  $\frac{\pi}{2} - \delta$ , gdzie  $\delta$  jest kątem między wektorami  $\mathbf{r}$  i  $\mathbf{v}$  (por. Rys. 1.2). A zatem

$$\begin{pmatrix} R \\ T \end{pmatrix} = \begin{pmatrix} \cos(\frac{\pi}{2} - \delta) & \sin(\frac{\pi}{2} - \delta) \\ -\sin(\frac{\pi}{2} - \delta) & \cos(\frac{\pi}{2} - \delta) \end{pmatrix} \begin{pmatrix} N \\ S \end{pmatrix}, \quad (1.66)$$

czyli

$$\begin{aligned} R &= N \sin \delta + S \cos \delta, \\ T &= -N \cos \delta + S \sin \delta, \end{aligned} \quad (1.67)$$

oraz

$$\begin{aligned} N &= R \sin \delta - T \cos \delta, \\ S &= R \cos \delta + T \sin \delta. \end{aligned} \quad (1.68)$$

Funkcje trygonometryczne kąta  $\delta$  są określone poprzez iloczyny wektorów  $\mathbf{r}$  i  $\mathbf{v}$

$$\begin{aligned} \cos \delta &= \frac{\mathbf{r} \cdot \mathbf{v}}{r v} = \frac{r \dot{r}}{r v} = \frac{\dot{r}}{v}, \\ \sin \delta &= \frac{|\mathbf{r} \times \mathbf{v}|}{r v} = \frac{G}{r v} = \frac{r \dot{f}}{v}. \end{aligned}$$

Jeśli sięgniemy do wzorów opisujących prędkość radialną  $\dot{r}$  i transwersalną  $r \dot{f}$  oraz przyjmiemy skrótowe oznaczenie

$$\gamma = \sqrt{1 + 2e \cos f + e^2}$$

tak, aby  $v = \gamma \sqrt{\mu/p}$ , to otrzymujemy

$$\begin{aligned} \cos \delta &= \frac{e \sin f}{\gamma}, \\ \sin \delta &= \frac{1 + e \cos f}{\gamma} = \frac{p}{r \gamma}. \end{aligned} \quad (1.69)$$

Baza perycentryczna spotykana jest rzadziej. Tworzą ją wersory:

- perycentryczny  $\hat{e}$ , czyli wersor wektora Laplace'a, skierowany do perycentrum,
- prostopadły do perycentrum  $\hat{Q} = \hat{b} \times \hat{e}$ , leżący w płaszczyźnie orbity (zwany też wersorem Hamiltona),
- binormalny  $\hat{b}$ .

W odróżnieniu od radialnej i stycznej, baza ta jest nieokreślona w przypadku orbit kołowych.

### 1.5.9 Ostateczna postać wzorów Gaussa

Po zdefiniowaniu baz, możemy przejść do ostatecznego uporządkowania równań Gaussa dla elementów oskulacyjnych wypisując każde z nich w trzech postaciach: najpierw w postaci wektorowej a następnie jako funkcje składowych  $R, T, B$  oraz  $N, S, B$ . Wyprowadzenie tych wzorów na podstawie równań (1.44, 1.45, 1.51, 1.52, 1.54, 1.63) jest elementarne i pozostawimy je jako samodzielne ćwiczenie. Podamy tylko wersję dla zaburzonych orbit eliptycznych.

Użyjemy oznaczeń  $s = \sin I$ ,  $c = \cos I$ ,  $p = a(1 - e^2)$ , oraz  $n = \sqrt{\mu a^{-3}}$ .  
Równania dla pięciu elementów oskulacyjnych to:

$$\dot{a} = \frac{2a^2}{\mu} \mathbf{P} \cdot \mathbf{v} = \frac{2}{n\sqrt{1-e^2}} \left[ R e \sin f + T \frac{p}{r} \right] = \frac{2v}{n^2 a} S, \quad (1.70)$$

$$\begin{aligned} \dot{e} &= \frac{1}{\mu a e} [(\mathbf{r} \cdot \mathbf{v}) \mathbf{r} + (ap - r^2) \mathbf{v}] \cdot \mathbf{P} \\ &= \frac{\sqrt{1-e^2}}{na} [R \sin f + T(\cos f + \cos E)] \\ &= \frac{1}{v} \left[ S \frac{2p}{r} \cos E + N \sqrt{1-e^2} \sin E \right], \end{aligned} \quad (1.71)$$

$$\dot{I} = \frac{r \cos(f + \omega)}{\mu p} (\mathbf{r} \times \mathbf{v}) \cdot \mathbf{P} = \frac{r \cos(f + \omega)}{\sqrt{\mu p}} B, \quad (1.72)$$

$$\begin{aligned} \dot{\omega} &= -\frac{r}{\mu p e} \left[ \frac{\sqrt{\mu p}}{r} (\cos E + e) \mathbf{r} - (p + r) \sin f \mathbf{v} \right] \cdot \mathbf{P} - c\dot{\Omega} \\ &= \frac{1}{e} \sqrt{\frac{p}{\mu}} \left[ -R \cos f + T \left( 1 + \frac{r}{p} \right) \sin f \right] - c\dot{\Omega} \\ &= \frac{1}{ev} [2S \sin f - N(e + \cos E)] - c\dot{\Omega}, \end{aligned} \quad (1.73)$$

$$\dot{\Omega} = \frac{r \sin(f + \omega)}{s \mu p} (\mathbf{r} \times \mathbf{v}) \cdot \mathbf{P} = \frac{r \sin(f + \omega)}{s \sqrt{\mu p}} B. \quad (1.74)$$

Do tego dochodzi szóste równanie dla anomalii średniej

$$\begin{aligned} \dot{M} &= n - \frac{1}{na^2} \left[ 2\mathbf{P} \cdot \mathbf{r} + \sqrt{\mu p} (\dot{\omega} + c\dot{\Omega}) \right] \\ &= n - \frac{1}{na^2 e} [R(2re - p \cos f) + T(r + p) \sin f] \\ &= n - \frac{\sqrt{1-e^2}}{ve} \left[ N(\cos f - e) - S \left( 1 + \frac{2re^2}{p} \right) \sin f \right], \end{aligned} \quad (1.75)$$

lub – jeszcze prostsze – dla anomalii prawdziwej

$$\dot{f} = \frac{\sqrt{\mu p}}{r^2} - \dot{\omega} - c\dot{\Omega}. \quad (1.76)$$

## WYKŁAD 6

### 1.5.10 Elementy nieosobliwe

Przyglądając się równaniom Gaussa zauważymy, że jeśli  $e = 0$ , to nie wolno używać wzorów dla  $\dot{e}$ ,  $\dot{M}$  i  $\dot{\omega}$ , gdyż albo wystąpi dzielenie przez zero, albo odwołamy się do nieokreślonych kątów  $f$  i  $E$ . Natomiast jeśli  $I = 0$  lub  $I = \pi$ , to nie wolno używać równań dla  $\dot{I}$  lub  $\dot{\Omega}$ , gdyż albo wystąpi dzielenie przez zero, albo pojawi się nieokreślony kąt  $f + \omega$ . Oczywiście, nie oznacza to, że nie potrafimy się obchodzić z takimi orbitami. Wystarczy wtedy stosować tzw. elementy nieosobliwe, będące funkcjami elementów keplerowskich. Są to np.

$$\begin{cases} q = \sin \frac{I}{2} \cos \Omega, \\ p = \sin \frac{I}{2} \sin \Omega, \\ \varpi = \omega + \Omega, \end{cases} \quad (1.77)$$

dla małych nachyleń lub

$$\begin{cases} k = e \cos \varpi, \\ h = e \sin \varpi, \\ \lambda = M + \varpi, \end{cases} \quad (1.78)$$

dla małych mimośrodków. Niestety, te tradycyjne symbole kolidują z równie tradycyjnym  $p$  – parametr orbity,  $q$  – odległość perycentrum,  $h$  – stała siły żywej,  $k$  – stała Gaussa...

Odpowiednia modyfikacja równań Gaussa prowadzi do równań pozbawionych osobliwości. Na przykład, dla małych mimośrodków używamy

$$\dot{\lambda} = \dot{M} + \dot{\varpi},$$

oraz

$$\begin{aligned} \dot{k} &= \dot{e} \cos \varpi - e \dot{\varpi} \sin \varpi = \frac{\dot{e} k}{\sqrt{h^2 + k^2}} - h \dot{\varpi}, \\ \dot{h} &= \dot{e} \sin \varpi + e \dot{\varpi} \cos \varpi = \frac{\dot{e} h}{\sqrt{h^2 + k^2}} + k \dot{\varpi}. \end{aligned}$$

Zauważmy, że zmienne  $h, k$  są poprawnie określone nie tylko dla orbit kołowych, lecz i dla orbit z zerowym nachyleniem. To, że kąt  $\varpi$  jako formalna suma  $\omega + \Omega$  nie traci sensu geometrycznego przy zerowym nachyleniu, jest oczywiste: przechodzi on gładko w kąt płaski między wersorami  $\hat{x}$  i  $\hat{e}$ .

Pokażmy jednak, że równanie Gaussa dla  $\dot{\varpi}$  nie zawiera osobliwości, jeśli wykonamy w nim stosowne przekształcenie. Przypomnijmy wzór (1.54), który przepiszemy w postaci

$$\dot{\omega} = \frac{F}{\mu e} - \cos I \dot{\Omega},$$

gdzie

$$F = 2(\mathbf{r} \cdot \hat{\mathbf{Q}})(\mathbf{v} \cdot \mathbf{P}) - (\mathbf{v} \cdot \hat{\mathbf{Q}})(\mathbf{r} \cdot \mathbf{P}) - (\mathbf{r} \cdot \mathbf{v})(\hat{\mathbf{Q}} \cdot \mathbf{P}).$$

Mamy więc dla długości perycentrum

$$\dot{\varpi} = \dot{\omega} + \dot{\Omega} = \frac{F}{\mu e} + (1 - \cos I) \dot{\Omega}.$$

W sytuacji zerowego nachylenia wyraz  $F/(\mu e)$  nie sprawia żadnych kłopotów, natomiast w  $\dot{\Omega}$ , które dane jest równaniem (1.74), pojawia się zrównoleżenie przez  $\sin I = 0$ , jak i nieokreślony kąt  $f + \omega$ :

$$\dot{\Omega} = \frac{r \sin(f + \omega)}{\sin I G} \mathbf{P} \cdot \hat{\mathbf{b}}.$$

Po wstawieniu (1.74), równanie dla  $\dot{\varpi}$  przybiera postać

$$\dot{\varpi} = \frac{F}{\mu e} + \left[ \frac{1 - \cos I}{\sin I} \right] \frac{r \sin(f + \omega)}{G} \mathbf{P} \cdot \hat{\mathbf{b}}.$$

Ale wyrażenie w nawiasie kwadratowym to nic innego jak tangens połowy nachylenia, więc mamy

$$\dot{\varpi} = \frac{F}{\mu e} + \operatorname{tg} \frac{I}{2} \frac{r \sin(f + \omega)}{G} \mathbf{P} \cdot \hat{\mathbf{b}}.$$

Przy nachyleniu  $I = 0$ , cały drugi składnik znika i problem nieokreślonego argumentu  $f + \omega$  przestaje być istotny. Oczywiście, bardziej eleganckim rozwiązaniem jest wyrugowanie kąta  $\omega$ , co łatwo zrealizować przy pomocy zmiennych  $p, q$ :

$$\begin{aligned} \sin(f + \omega) &= \sin(f + \varpi - \Omega) = \sin(f + \varpi) \cos \Omega - \cos(f + \varpi) \sin \Omega = \\ &= \frac{q \sin(f + \varpi) - p \cos(f + \varpi)}{\sin \frac{I}{2}}, \end{aligned}$$

i tak otrzymujemy jawnie nieosobliwe

$$\dot{\varpi} = \frac{F}{\mu e} + \frac{r(q \sin(f + \varpi) - p \cos(f + \varpi))}{G \sqrt{1 - p^2 - q^2}} \mathbf{P} \cdot \hat{\mathbf{b}}, \quad (1.79)$$

gdzie podstawiliśmy  $\cos I/2 = \sqrt{1 - \sin^2 I/2} = \sqrt{1 - (p^2 + q^2)}$ .

Przekształcanie do postaci nieosobliwej ze względu na mimośród przebiega podobnie. Odpowiednikiem wykorzystania tangensa połowy nachylenia jest wtedy przekształcenie

$$\frac{1 - \sqrt{1 - e^2}}{e} = \frac{e}{1 + \sqrt{1 - e^2}}.$$

## 1.6 Przykładowe zastosowania równań Gaussa

Spróbujemy teraz zilustrować zastosowania równań Gaussa zanim jeszcze zajmiemy się ogólnymi metodami rozwiązywania równań ruchu. Istnieją dwa przypadki, w których zastosowanie równań Gaussa staje się uzasadnione: po pierwsze – gdy mamy do czynienia z siłą dyssypatywną jak na przykład opór ośrodka, po drugie – gdy stosujemy tzw. *przybliżenie impulsowe*, w którym siła zmienia się skokowo (fizycznie odpowiada to takim efektom jak odpalenie silnika statku kosmicznego, zaburzenie ruchu przez pojedyncze i krótkotrwałe zbliżenie z trzecim ciałem itp.). W pierwszym przypadku nie można wprowadzić potencjału, co uniemożliwia zastosowanie formalizmu bardziej szczegółowego niż newtonowski. W drugim – posługiwanie się potencjałem jest w zasadzie możliwe, lecz zazwyczaj nieopłacalne (trzeba wprowadzić funkcje Diraca). Przedstawmy przykłady obu zastosowań.

### 1.6.1 Wpływ oporu ośrodka na elementy oskulacyjne.

Rozpatrzmy względne zagadnienie dwóch ciał, w którym na badaną masę działa hamująca siła  $\mathbf{P}$  wywołana oporem nieruchomego ośrodka. Siła tego rodzaju skierowana jest stycznie do chwilowej orbity, w kierunku przeciwnym do wektora prędkości. Możemy więc zauważyć, że jej składowe w bazie stycznej mają wartości

$$S = S(\mathbf{r}, \mathbf{v}) < 0, \quad N = B = 0,$$

gdzie bez względu na postać zależności tarcia od położenia ciała  $\mathbf{r}$  i jego prędkości  $\mathbf{v}$ , poprzestajemy jedynie na stwierdzeniu, że  $S < 0$  a więc opór nie zanika ani (tym bardziej) nie zmienia znaku. Przyjrzyjmy się równaniom Gaussa (1.70–1.74), w których jedyną niezerową składową będzie  $S$

$$\begin{aligned} \dot{a} &= \frac{2v}{n^2 a} S < 0, \\ \dot{e} &= \frac{2pS}{rv} \cos E, \end{aligned}$$

$$\begin{aligned} \dot{I} &= \dot{\Omega} = 0, \\ \dot{\omega} &= \frac{2S}{ev} \sin f. \end{aligned} \tag{1.80}$$

Pomijając równanie dla anomalii średniej (1.75) jako mniej istotne, zobaczymy jak wiele ważnych wniosków o ewolucji orbity można wysnuć z samych tylko równań Gaussa (bez ich rozwiązywania):

1. Półoś wielka orbity systematycznie maleje, gdyż  $\dot{a} < 0$ . W konsekwencji może nastąpić kolizja obu mas, ale pod warunkiem, że spadek półosi do zera wymaga skończonego czasu.
2. Działająca stycznie siła tarcia nie powoduje zmian w orientacji płaszczyzny orbity, gdyż nachylenie  $I$  oraz długość linii węzłów  $\Omega$  pozostają stałe. A zatem trajektoria w tym zagadnieniu jest krzywą płaską o kształcie spiralnym (malejące  $a$ ).
3. Zmiany oskulacyjnego mimośrod  $e$  zależą w sposób istotny od położenia badanej masy na orbicie: w pobliżu perycentrum ( $\cos E > 0$ ) opór ośrodka zmniejsza mimośród ( $\dot{e} < 0$ ) natomiast w pobliżu apocentrum ( $\cos E < 0$ ) mimośród rośnie ( $\dot{e} > 0$ ).

Powstaje w tym momencie interesujące pytanie: czy zmiany mimośrodu mają charakter czysto okresowy? Innymi słowy, czy podczas każdego obiegu przyrost mimośrodu w pobliżu apocentrum jest w pełni kompensowany przez jego spadek w pobliżu perycentrum? Taka pełna symetria efektu zapobiegłaby systematycznym zmianom mimośrodu. Dokładna odpowiedź na to pytanie wymaga nie tylko znajomości dokładnej postaci siły  $\mathbf{P}$  lecz również zastosowania odpowiedniej metody rozwiązywania równań ruchu. Można jednak pokusić się o pewne oszacowania: uwzględniając fizykę zagadnienia można przyjąć, że  $S(\mathbf{r}, \mathbf{v})$  jest proporcjonalna do dodatniej potęgi  $v$  (najczęściej jest to  $v^1$ ,  $v^{\frac{3}{2}}$  lub  $v^2$ ) oraz do niedodatniej potęgi odległości  $r^\beta$ , gdzie  $\beta \leq 0$  (tzn. gęstość ośrodka jest albo stała albo maleje w miarę oddalania się od masy centralnej).

Założmy, dla uproszczenia, że w ciągu jednego obiegu obowiązuje całka pól, a więc wartości iloczynu  $rv$  w perycentrum i w apocentrum są sobie równe (tzw. zasada dźwigni). W takim razie, oznaczając  $\alpha \equiv 2p/(rv)$ , mamy w perycentrum

$$\dot{e}_P = \alpha S(q, v_P) < 0,$$

a w apocentrum

$$\dot{e}_A = -\alpha S(Q, v_A) > 0.$$



Zauważmy teraz, że  $|S(q, v_P)| > |S(Q, v_A)|$ , co sugerują zarówno prędkości  $v_P > v_A$  jak i odległości  $q < Q$ . Jak widać, **opór ośrodka powoduje systematyczny spadek mimośrodu orbity**, gdyż  $\dot{e}_P + \dot{e}_A < 0$ .

Przedstawione tu rozumowanie wymaga dwóch istotnych zastrzeżeń: po pierwsze, oparte jest na porównaniu dwóch punktów orbity i używa dość grubych przybliżeń; po drugie – nie można twierdzić, jakoby mimośród asymptotycznie dążył do zera, gdyż równania Gaussa dla  $\dot{e}$  nie wolno stosować w przypadku małych mimośrodów. Malejący systematycznie mimośród jest typową cechą orbit eliptycznych poddanych oporowi ośrodka, ale wnioski o charakterze asymptotycznym, typu  $\lim_{t \rightarrow \infty} e = 0$ , często są wyciągane zbyt pochopnie.

### 1.6.2 Manewry orbitalne (zmiana nachylenia)

Jednym z prostszych zastosowań równań Gaussa jest opis manewru orbitalnego w przybliżeniu impulsowym. Manewr orbitalny to zmiana elementów orbity dokonywana przy pomocy silników korekcyjnych sztucznego satelity lub sondy kosmicznej. Przybliżenie impulsowe natomiast zakłada, że silniki działają stałą siłą  $\mathbf{P}$  (zwaną **ciągiem**) przez krótki czas  $\Delta t$  i zmiana w dowolnym elemencie orbity (na przykład w półośi wielkiej  $a$ ) może być przedstawiona jako

$$\Delta a \approx \dot{a} \Delta t. \quad (1.81)$$

W każdym manewrze orbitalnym dąży się do uzyskania zamierzonej orbity przy minimalnym wydatku paliwa, co oznacza, że staramy się użyć jak najmniejszego **impulsu** (zwanego również popędem). W naszych rozważaniach będziemy zajmować się impulsem podzielonym przez masę, czyli

$$W = |\mathbf{P}| \Delta t, \quad (1.82)$$

gdź  $\mathbf{P}$  jest przyspieszeniem, a nie siłą.

Dla zilustrowania problemu zajmiemy się manewrem polegającym na zmianie nachylenia orbity.

W świetle równania Gaussa (1.72) jedynie składowa normalna do płaszczyzny orbity powoduje zmianę nachylenia  $I$ , a zatem siła powodująca manewr powinna być skierowana prostopadle do płaszczyzny orbity i odpowiada jej impuls

$$W = B \Delta t.$$

Uzyskany przyrost nachylenia dany będzie wzorem

$$\Delta I = \frac{r \cos(f + \omega)}{\sqrt{\mu p}} W. \quad (1.83)$$

Zastanówmy się teraz, w którym punkcie orbity należy uruchomić silniki aby uzyskać największy przyrost  $\Delta I$  przy zadanym  $W$ . Jedyną wielkością we wzorze (1.83) co do której mamy jakiś wybór jest anomalia prawdziwa  $f$  a więc musimy znaleźć maksimum funkcji  $r \cos(f + \omega)$  z warunku

$$\frac{d(r \cos(f + \omega))}{df} = 0.$$

Uwzględniając zależność  $r(f)$  otrzymamy

$$-p \frac{\sin(f + \omega)(1 + e \cos f) - e \sin f \cos(f + \omega)}{(1 + e \cos f)^2} = 0,$$

a po odrzuceniu niezerowego czynnika  $p/(1 + e \cos f)^2$  i skorzystaniu z wzoru dla sinusa różnicy kątów dochodzimy ostatecznie do warunku

$$\sin(f + \omega) + e \sin \omega = 0. \quad (1.84)$$

Dla małych wartości mimośrodoru orbity  $e$  równanie (1.84) prowadzi do wniosku, że manewr zmiany nachylenia należy przeprowadzić w okolicach jednego z węzłów orbity, gdy  $(f + \omega) \approx 0$  lub  $\pi$ . Analizując pozostałe równania Gaussa dochodzimy do wniosku, że manewr impulsem normalnym do płaszczyzny orbity przeprowadzony dokładnie w węźle spowoduje **jedynie** zmianę nachylenia nie wpływając na pozostałe elementy orbity.

Zauważmy również, że przyrost  $I$  jest wprost proporcjonalny do odległości  $r$  („ramię siły”), a zatem szczególnie uprzywilejowana będzie sytuacja, w której linia apsyd i linia węzłów pokrywają się, aby manewr można przeprowadzić zarazem w węźle i w apocentrum.

## Rozdział 2

# Formalizm kanoniczny

### WYKŁAD 7

Formalizm kanoniczny nazywany jest również formalizmem symplektycznym i stanowi on szczególnie przypadek formalizmu Hamiltona. Podstawowe pojęcia tego formalizmu to położenia i pędy uogólnione oraz funkcja Hamiltona. Ruch układu mechanicznego zadany jest poprzez równania kanoniczne Hamiltona. W odróżnieniu od formalizmu newtonowskiego, zakres stosowalności formalizmu kanonicznego jest ograniczony, co bierze się z faktu, że cała dynamika układu jest „zaszyfrowana” w jednej skalarnej funkcji stanu (funkcji Hamiltona), która generuje prawe strony równań ruchu poprzez swój gradient. Musimy więc ograniczyć się do zagadnień z siłami potencjalnymi. Za tę cenę nabywamy jednak możliwość względnie prostego formułowania twierdzeń opisujących całe klasy ruchów.

Z wielu sposobów wprowadzania formalizmu kanonicznego wybierzmy teorię form i równań Johanna Friedricha Pfaffa (1815). Najpierw jednak wprowadzimy kilka pojęć związanych z grupą macierzy symplektycznych.

## 2.1 Macierze i formy symplektyczne

### 2.1.1 Standardowa macierz symplektyczna i forma symplektyczna

Oznaczmy przez  $\mathbf{J}$  **standardową macierz symplektyczną**. Jest to antysymetryczna macierz kwadratowa stopnia parzystego  $2M \times 2M$ , którą można zapisać w postaci blokowej jako

$$\mathbf{J} = \begin{pmatrix} \mathbf{0}_M & \mathbf{E}_M \\ -\mathbf{E}_M & \mathbf{0}_M \end{pmatrix}. \quad (2.1)$$

Ponieważ  $\mathbf{E}_M$  oznacza macierz jednostkową  $M \times M$ , to – na przykład – dla dwóch stopni swobody mamy

$$\mathbf{J} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}.$$

Często można się spotkać z definicją standardowej macierzy symplektycznej która ma w porównaniu z (2.1) przestawione znaki, ale nie ma to wpływu na większość z dalszych wzorów.

Standardowa macierz symplektyczna jest obiektem posiadającym szereg interesujących właściwości:

$$\mathbf{J}^2 \equiv \mathbf{J}\mathbf{J} = -\mathbf{E}_{2M}, \quad \mathbf{J}^T\mathbf{J} = \mathbf{E}_{2M}, \quad \mathbf{J}^{-1} = \mathbf{J}^T = -\mathbf{J}, \quad \det \mathbf{J} = 1. \quad (2.2)$$

Z dwóch wektorów  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{2M}$  i macierzy  $\mathbf{J}$  stopnia  $2M$  można utworzyć algebraiczną formę dwuliniową<sup>1</sup>

$$[\mathbf{x}|\mathbf{y}] = \mathbf{x}^T\mathbf{J}\mathbf{y} = \mathbf{x} \cdot (\mathbf{J}\mathbf{y}), \quad (2.3)$$

która nosi nazwę **formy symplektycznej**. Jest ona antysymetryczna, gdyż

$$[\mathbf{y}|\mathbf{x}] = \mathbf{y} \cdot (\mathbf{J}\mathbf{x}) = (\mathbf{J}\mathbf{x}) \cdot \mathbf{y} = (\mathbf{J}\mathbf{x})^T \mathbf{y} = \mathbf{x}^T \mathbf{J}^T \mathbf{y} = -\mathbf{x}^T \mathbf{J}\mathbf{y} = -[\mathbf{x}|\mathbf{y}],$$

gdzie skorzystaliśmy z przemienności iloczynu skalarnego i antysymetryczności macierzy  $\mathbf{J}$ . Jak zwykle, antysymetryczność oznacza  $[\mathbf{x}|\mathbf{x}] = 0$ .

Jeśli w wektorach  $2M$ -wymiarowych wydzielimy dwa  $M$ -wymiarowe bloki, na przykład

$$\mathbf{x} = \text{col}(\mathbf{q}, \mathbf{Q}), \quad \mathbf{y} = \text{col}(\mathbf{p}, \mathbf{P}),$$

to forma symplektyczna jest różnicą iloczynów skalarnych

$$[\mathbf{x}|\mathbf{y}] = \mathbf{q} \cdot \mathbf{P} - \mathbf{p} \cdot \mathbf{Q} = \mathbf{q}^T \mathbf{P} - \mathbf{p}^T \mathbf{Q}. \quad (2.4)$$

### 2.1.2 Definicja macierzy symplektycznych

Pamiętamy, że przekształcenia ortogonalne

$$\mathbf{x} \mapsto \mathbf{x}' = \mathbf{M}\mathbf{x}, \quad \mathbf{y} \mapsto \mathbf{y}' = \mathbf{M}\mathbf{y},$$

---

<sup>1</sup>Forma dwuliniowa to funkcja, która dwóm wektorom przyporządkowuje skalar i jest liniowa względem obu argumentów. Najbardziej znany przykład to iloczyn skalarny wektorów.

zachowują iloczyn skalarny, gdyż

$$\mathbf{x}' \cdot \mathbf{y}' = (\mathbf{M}\mathbf{x}) \cdot (\mathbf{M}\mathbf{y}) = \mathbf{x}^T \mathbf{M}^T \mathbf{M} \mathbf{y} = \mathbf{x} \cdot \mathbf{y}.$$

W podobny sposób możemy zdefiniować liniowe przekształcenia symplektyczne

$$\mathbf{x} \mapsto \mathbf{x}' = \mathbf{S}\mathbf{x}, \quad \mathbf{y} \mapsto \mathbf{y}' = \mathbf{S}\mathbf{y},$$

postulując, aby dla każdej pary  $\mathbf{x} \neq \mathbf{y}$  zachowywały one wartość formy symplektycznej

$$[\mathbf{x}' | \mathbf{y}'] = [\mathbf{S}\mathbf{x} | \mathbf{S}\mathbf{y}] = [\mathbf{x} | \mathbf{y}]. \quad (2.5)$$

To oznacza warunek

$$\mathbf{S}^T \mathbf{J} \mathbf{S} = \mathbf{J}, \quad (2.6)$$

i każdą macierz  $\mathbf{S}$ , która go spełnia, nazywamy **macierzą symplektyczną**.

Korzystając z własności (2.2) łatwo można sprawdzić, że macierz standardowa  $\mathbf{J}$  jest macierzą symplektyczną, gdyż

$$\mathbf{J}^T \mathbf{J} \mathbf{J} = \mathbf{E} \mathbf{J} = \mathbf{J}$$

(od tego momentu opuszczamy indeks  $2M$  macierzy jednostkowej dla skrócenia zapisu). Także macierz jednostkowa  $\mathbf{E}$  jest macierzą symplektyczną:

$$\mathbf{E}^T \mathbf{J} \mathbf{E} = \mathbf{E} \mathbf{J} \mathbf{E} = \mathbf{J} \mathbf{E} = \mathbf{J}.$$

### 2.1.3 Grupa symplektyczna

Pokażemy teraz, że macierze symplektyczne stopnia  $2M$  z działaniem mnożenia macierzy tworzą grupę, nazywaną **grupą symplektyczną**  $Sp(2M)$ .

Najpierw musimy wykazać, że iloczyn dwóch macierzy symplektycznych daje macierz symplektyczną, czyli że mnożenie jest działaniem wewnętrznym w  $Sp(2M)$ . Niech  $\mathbf{S}_1, \mathbf{S}_2 \in Sp(2M)$ . Wtedy, wychodząc od macierzy  $\mathbf{J}$ , możemy zastosować dwa razy definicję (2.6)

$$\mathbf{J} = \mathbf{S}_2^T \mathbf{J} \mathbf{S}_2 = \mathbf{S}_2^T (\mathbf{S}_1^T \mathbf{J} \mathbf{S}_1) \mathbf{S}_2 = (\mathbf{S}_2^T \mathbf{S}_1^T) \mathbf{J} (\mathbf{S}_1 \mathbf{S}_2) = (\mathbf{S}_1 \mathbf{S}_2)^T \mathbf{J} (\mathbf{S}_1 \mathbf{S}_2),$$

a to znaczy, że  $\mathbf{S} = \mathbf{S}_1 \mathbf{S}_2$  jest macierzą symplektyczną.

Pozostały do sprawdzenia trzy warunki: łączność mnożenia, istnienie elementu neutralnego w zbiorze macierzy symplektycznych oraz istnienie elementu odwrotnego dla każdej macierzy symplektycznej. Pierwsze dwa są spełnione, gdyż łączność jest podstawową cechą iloczynu macierzy, zaś jego element neutralny, macierz  $\mathbf{E}$ , jest macierzą symplektyczną, co pokazaliśmy

wyżej. Pozostaje więc problem: czy każda macierz symplektyczna  $\mathbf{S}$  posiada macierz odwrotną  $\mathbf{S}^{-1}$ , a jeśli tak, to czy  $\mathbf{S}^{-1}$  jest macierzą symplektyczną?

O odwracalności macierzy decyduje wartość wyznacznika, która nie może być równa 0. Można to sprawdzić w prosty sposób. Wykorzystujemy definicję (2.6) oraz twierdzenie, w myśl którego wyznacznik iloczynu macierzy jest równy iloczynowi ich wyznaczników

$$\det(\mathbf{S}^T \mathbf{J} \mathbf{S}) = \det \mathbf{J} \Rightarrow \det \mathbf{S}^T \det \mathbf{J} \det \mathbf{S} = \det \mathbf{J}.$$

Ponieważ  $\det \mathbf{J} = 1$  i  $\det \mathbf{S}^T = \det \mathbf{S}$ , to

$$(\det \mathbf{S})^2 = 1 \neq 0.$$

Co więcej, można udowodnić, że

$$\det \mathbf{S} = 1. \quad (2.7)$$

Niestety – dowód, że  $\det \mathbf{S} \neq -1$  nie należy do trywialnych i pominiemy go.

Skoro wyznacznik macierzy symplektycznej jest różny od zera, to każda macierz symplektyczna posiada macierz odwrotną  $\mathbf{S}^{-1}$ , dla której

$$\mathbf{S}^{-1} \mathbf{S} = \mathbf{S} \mathbf{S}^{-1} = \mathbf{E}.$$

Czy  $\mathbf{S}^{-1}$  jest również macierzą symplektyczną? Łatwo to sprawdzić:

$$\mathbf{J} = \mathbf{E}^T \mathbf{J} \mathbf{E} = (\mathbf{S} \mathbf{S}^{-1})^T \mathbf{J} (\mathbf{S} \mathbf{S}^{-1}) = (\mathbf{S}^{-1})^T (\mathbf{S}^T \mathbf{J} \mathbf{S}) (\mathbf{S}^{-1}).$$

A ponieważ  $\mathbf{S}$  jest symplektyczna, to środkowy iloczyn wynosi  $\mathbf{J}$  i

$$\mathbf{J} = (\mathbf{S}^{-1})^T \mathbf{J} \mathbf{S}^{-1},$$

co w świetle definicji (2.6) oznacza, że  $\mathbf{S}^{-1}$  jest macierzą symplektyczną.

Okazuje się przy tym, że związek między  $\mathbf{S}^{-1}$  i  $\mathbf{S}$  jest niemal równie prosty, jak w wypadku macierzy ortogonalnych, gdzie  $\mathbf{M}^{-1} = \mathbf{M}^T$ .

$$\begin{aligned} \mathbf{S} \mathbf{S}^{-1} &= \mathbf{E}, & / \mathbf{S}^T \mathbf{J} \times \\ (\mathbf{S}^T \mathbf{J} \mathbf{S}) \mathbf{S}^{-1} &= \mathbf{S}^T \mathbf{J}, \\ \mathbf{J} \mathbf{S}^{-1} &= \mathbf{S}^T \mathbf{J}, & / \mathbf{J}^T \times \\ \mathbf{J}^T \mathbf{J} \mathbf{S}^{-1} &= \mathbf{J}^T \mathbf{S}^T \mathbf{J}, \end{aligned}$$

Korzystając z (2.2) dochodzimy do

$$\mathbf{S}^{-1} = \mathbf{J}^T \mathbf{S}^T \mathbf{J}. \quad (2.8)$$

W grupie symplektycznej mamy jeszcze jedną interesującą własność:  $\mathbf{S}^T \in Sp(2M)$ . Zaczniemy od

$$\mathbf{S}\mathbf{S}^{-1} = \mathbf{E}.$$

Podstawiamy (2.8), co daje

$$\mathbf{S}\mathbf{J}^T\mathbf{S}^T\mathbf{J} = \mathbf{E}.$$

Mnożymy prawostronnie przez  $\mathbf{J}$

$$\mathbf{S}\mathbf{J}^T\mathbf{S}^T\mathbf{J}\mathbf{J} = \mathbf{J}.$$

i uwzględniając  $\mathbf{J}\mathbf{J} = -\mathbf{E}$  oraz  $\mathbf{J}^T = -\mathbf{J}$  z własności (2.2), dochodzimy do

$$\mathbf{S}\mathbf{J}\mathbf{S}^T = \mathbf{J}. \tag{2.9}$$

To zaś jest równoważne stwierdzeniu, że  $\mathbf{S}^T$  spełnia warunek (2.6), gdyż

$$\mathbf{S}\mathbf{J}\mathbf{S}^T = (\mathbf{S}^T)^T \mathbf{J} (\mathbf{S}^T) = \mathbf{J}.$$

## 2.2 Pfaffian i równania Pfaffa

Wprowadźmy wektor  $N + 1$  zmiennych

$$\mathbf{w} = (w_0, w_1, \dots, w_N)^T \in \mathbb{R}^{N+1},$$

oraz wektor  $N + 1$  funkcji tych zmiennych

$$\mathbf{F}(\mathbf{w}) = (F_0(\mathbf{w}), F_1(\mathbf{w}), \dots, F_N(\mathbf{w}))^T \in \mathbb{R}^{N+1}.$$

Z tych dwóch elementów możemy utworzyć liniową **formę różniczkową Pfaffa** (a ściślej, tzw. jednoformę różniczkową), czyli **pfaffian**

$$\Phi = \sum_{i=0}^N F_i dw_i = \mathbf{F} \cdot d\mathbf{w} = \mathbf{F}^T d\mathbf{w}, \quad (2.10)$$

gdzie  $d\mathbf{w}$  jest wektorem różniczek zmiennych  $w_i$ .

Forma Pfaffa sama w sobie jest tylko pewnym wyrażeniem matematycznym i można ją wykorzystać do różnych celów. Nas interesować będą równania związane z  $\Phi$ , zwane **stowarzyszonym układem równań Pfaffa** (pierwszego rodzaju)

$$\left[ D\mathbf{w}\mathbf{F} - (D\mathbf{w}\mathbf{F})^T \right] d\mathbf{w} = \mathbf{0}, \quad (2.11)$$

gdzie transponowany gradient zastosowany do kolejnych elementów wektora kolumnowego  $\mathbf{F}$  daje macierz Jacobiego

$$D\mathbf{w}\mathbf{F} = \mathbf{A}, \quad a_{ij} = \frac{\partial F_i}{\partial w_j},$$

w której wiersze  $i$  oraz kolumny  $j$  indeksujemy od 0 do  $N$ . Wyrażenie w nawiasie kwadratowym jest macierzą antysymetryczną o wymiarze  $(N+1) \times (N+1)$  (tensorowym uogólnieniem operatora rotacji) zawierającą pochodne cząstkowe funkcji  $F_i$  z formy Pfaffa, więc (2.11) jest równaniem wektorowym, które można rozbić na  $N + 1$  równań skalarnych

$$\sum_{j=0}^N \left( \frac{\partial F_i}{\partial w_j} - \frac{\partial F_j}{\partial w_i} \right) dw_j = 0, \quad (2.12)$$

dla kolejnych  $i = 0, \dots, N$ .

Układ równań Pfaffa posiada ważną własność:



**TWIERDZENIE 2** Jeśli  $\Phi = \mathbf{F}(\mathbf{w}) \cdot d\mathbf{w}$  jest formą Pfaffa,  $c \neq 0$  dowolną stałą, zaś  $S(\mathbf{w})$  dowolną skalarną funkcją różniczkowalną, to forma  $\Phi$  oraz forma  $\Phi' = c(\Phi + dS(\mathbf{w}))$  generują ten sam układ równań Pfaffa.

**DOWÓD:** Równania generowane przez  $\Phi$  są dane wzorem (2.12). Musimy wykazać, że są one identyczne, jeśli użyjemy formy

$$\Phi' = c(\Phi + dS(\mathbf{w})) = c(\mathbf{F} \cdot d\mathbf{w} + dS(\mathbf{w})).$$

Różniczka zupełna definiowana jest jako

$$dS = \sum_{j=0}^N \frac{\partial S}{\partial w_j} dw_j = \nabla_{\mathbf{w}} S \cdot d\mathbf{w} = D_{\mathbf{w}} S d\mathbf{w},$$

więc

$$\Phi' = c\mathbf{F}(\mathbf{w}) \cdot d\mathbf{w} + c(\nabla_{\mathbf{w}} S) \cdot d\mathbf{w} = \mathbf{F}' \cdot d\mathbf{w}.$$

gdzie

$$\mathbf{F}' = c(\mathbf{F} + \nabla_{\mathbf{w}} S).$$

Równania Pfaffa otrzymane z  $\Phi'$  mają postać

$$\sum_{j=0}^N \left( \frac{\partial F'_i}{\partial w_j} - \frac{\partial F'_j}{\partial w_i} \right) dw_j = 0,$$

gdzie

$$F'_i = cF_i + c \frac{\partial S}{\partial w_i}, \quad F'_j = cF_j + c \frac{\partial S}{\partial w_j},$$

czyli

$$\sum_{j=0}^N \left( c \frac{\partial F_i}{\partial w_j} + c \frac{\partial^2 S}{\partial w_i \partial w_j} - c \frac{\partial F_j}{\partial w_i} - c \frac{\partial^2 S}{\partial w_j \partial w_i} \right) dw_j = 0.$$

Drugie pochodne  $S$  odejmują się, zaś obie strony można podzielić przez niezerową stałą  $c$ , co prowadzi do postaci (2.12).

Możemy więc przyjąć, że z punktu widzenia równań Pfaffa forma  $\Phi$  posiada tak zwaną swobodę cechowania (ang. *gauge freedom*), gdyż można do niej dodać dowolną różniczkę zupełną bez wpływu na generowane równania. Także swoboda pomnożenia pfaffianu przez dowolną stałą jest zaletą nie do

pogardzenia.

Jeśli jedną ze zmiennych w wektorze  $\mathbf{w}$ , na przykład  $w_0$ , wybierzemy jako zmienną niezależną, to dzieląc równania Pfaffa przez  $dw_0$  sprowadzimy je do układu równań różniczkowych zwyczajnych

$$\sum_{j=0}^N \left( \frac{\partial F_j}{\partial w_i} - \frac{\partial F_i}{\partial w_j} \right) \frac{dw_j}{dw_0} = 0,$$

dla  $N$  zmiennych zależnych  $w_1(w_0), \dots, w_N(w_0)$ . W ogólnym przypadku jest to postać niezbyt atrakcyjna, daleka od standardowej postaci typu (1.3). Są jednak szczególne przypadki, w których powstawać będą stosunkowo proste układy równań różniczkowych zwyczajnych – od razu w postaci standardowej.

Ten pozornie prosty zabieg dzielenia powoduje jednak radykalną zmianę perspektywy: pojawienie się pochodnych  $\frac{dw_j}{dw_0}$  oznacza, że zaczynamy rozumieć  $w_j$  nie jako niezależne zmienne, ale jako funkcje zmiennej  $w_0$ , których wartościami  $w_j(w_0)$  są punkty leżące na pewnej krzywej w  $N$ -wymiarowej przestrzeni  $(w_1, \dots, w_N)$ .

### 2.3 Równania kanoniczne jako przypadek szczególny równań Pfaffa

Przyjmijmy, że  $N = 2M$  i weźmy następujące wektory  $\mathbf{w}, \mathbf{F} \in \mathbb{R}^{2M+1}$

$$\begin{aligned} \mathbf{w} &= \text{col}(t, \mathbf{q}, \mathbf{Q}) = \text{col}(t, \boldsymbol{\xi}), \\ \mathbf{F} &= \text{col} \left( \mathcal{H}(\mathbf{w}), -\frac{\mathbf{Q}}{2}, \frac{\mathbf{q}}{2} \right) = \text{col} \left( \mathcal{H}(\boldsymbol{\xi}, t), -\frac{\mathbf{J}\boldsymbol{\xi}}{2} \right), \end{aligned} \quad (2.13)$$

gdzie  $t, \mathcal{H} \in \mathbb{R}$ ,  $\mathbf{q}, \mathbf{Q} \in \mathbb{R}^M$ , oraz  $\boldsymbol{\xi} \in \mathbb{R}^{2M}$ , co oznacza, że używamy standardowej macierzy symplektycznej  $\mathbf{J}$  stopnia  $2M$ . Nadajmy teraz nazwy poszczególnym klockom, których użyliśmy w (2.13):

- Wektor  $\mathbf{q} = (q_1, \dots, q_M)^T$  nazwiemy **współrzędnymi uogólnionymi**. Należy on do  $M$ -wymiarowej **przestrzeni konfiguracyjnej**.
- Wektor  $\mathbf{Q} = (Q_1, \dots, Q_M)^T$  nazwiemy **pędami uogólnionymi**. Należy on do  $M$ -wymiarowej **przestrzeni pędów**.

- O parze zmiennych  $q_j$  oraz  $Q_j$  (z tym samym indeksem) mówimy, że są kanonicznie sprzężone.
- Wektor  $\boldsymbol{\xi} = \text{col}(\mathbf{q}, \mathbf{Q})$  nazwiemy **wektorem stanu**. Należy on do  $2M$ -wymiarowej **przestrzeni fazowej**.
- Funkcję skalarną  $\mathcal{H}(\mathbf{w}) = \mathcal{H}(\mathbf{q}, \mathbf{Q}, t)$  nazwiemy **funkcją Hamiltona** lub **hamiltonianem**.

Zmienną  $t$  wybieramy jako zmienną niezależną i tak też będziemy ją na razie nazywać.

Wektory (2.13) definiują formę Pfaffa

$$\Phi = \mathcal{H} dt - \frac{\mathbf{Q} \cdot d\mathbf{q}}{2} + \frac{\mathbf{q} \cdot d\mathbf{Q}}{2}, \quad (2.14)$$

w której rozpoznajemy formę symplektyczną (2.4)

$$\Phi = \mathcal{H} dt + \frac{[\boldsymbol{\xi} | d\boldsymbol{\xi}]}{2}. \quad (2.15)$$

Jak wyglądają równania generowane przez ten pfaffian? Posłużymy się postacią wektorowo-macierzową (2.11), ale ponieważ  $\mathbf{w} = \text{col}(t, \boldsymbol{\xi})$ , rozbijemy operator  $D\mathbf{w}$  na dwie części:

$$D\mathbf{w} = \left( D_t, D_{\boldsymbol{\xi}} \right),$$

w postaci blokowej, gdzie  $D_t$  to zwykła pochodna cząstkowa  $\frac{\partial}{\partial t}$ . A zatem, działając na wektor  $\mathbf{F}$  dany wzorem (2.13), mamy

$$D\mathbf{w}\mathbf{F} = \begin{pmatrix} D_t\mathcal{H} & D_{\boldsymbol{\xi}}\mathcal{H} \\ D_t(-\mathbf{J}\boldsymbol{\xi}/2) & D_{\boldsymbol{\xi}}(-\mathbf{J}\boldsymbol{\xi}/2) \end{pmatrix}.$$

Przypomnijmy, że  $D_{\boldsymbol{\xi}}\boldsymbol{\xi} = \mathbf{E}$ , więc

$$D_{\boldsymbol{\xi}}(-\mathbf{J}\boldsymbol{\xi}) = -\mathbf{J} D_{\boldsymbol{\xi}}\boldsymbol{\xi} = -\mathbf{J}.$$

Ponadto, zmienne  $\mathbf{w}$  traktujemy jako niezależne, zatem  $D_t\boldsymbol{\xi} = \mathbf{0}$ . Możemy więc uprościć macierz Jacobiego  $D\mathbf{w}\mathbf{F}$  do postaci

$$D\mathbf{w}\mathbf{F} = \begin{pmatrix} D_t\mathcal{H} & D_{\boldsymbol{\xi}}\mathcal{H} \\ \mathbf{0} & -\mathbf{J}/2 \end{pmatrix}$$

a jej transpozycję do

$$(D\mathbf{w}\mathbf{F})^T = \begin{pmatrix} D_t\mathcal{H} & \mathbf{0} \\ \nabla_{\boldsymbol{\xi}}\mathcal{H} & \mathbf{J}/2 \end{pmatrix},$$

gdzie skorzystaliśmy z  $D_{\boldsymbol{\xi}} = \nabla_{\boldsymbol{\xi}}^T$ , oraz  $\mathbf{J}^T = -\mathbf{J}$ .

Równania Pfaffa (2.11) przybierają więc w obecnie rozpatrywanym przypadku postać

$$\begin{pmatrix} 0 & D_{\boldsymbol{\xi}}\mathcal{H} \\ -\nabla_{\boldsymbol{\xi}}\mathcal{H} & -\mathbf{J} \end{pmatrix} \begin{pmatrix} dt \\ d\boldsymbol{\xi} \end{pmatrix} = \mathbf{0}.$$

Wypisując osobno górny i dolny blok, mamy

$$(D_{\boldsymbol{\xi}}\mathcal{H}) d\boldsymbol{\xi} = 0 \quad (2.16)$$

$$-(\nabla_{\boldsymbol{\xi}}\mathcal{H}) dt - \mathbf{J}d\boldsymbol{\xi} = \mathbf{0}. \quad (2.17)$$

Przyjrzyjmy się równaniom (2.17). Dzieląc stronami przez  $dt$  i mnożąc lewostronnie przez  $\mathbf{J}$  otrzymujemy

$$\frac{d\boldsymbol{\xi}}{dt} = \dot{\boldsymbol{\xi}} = \mathbf{J}\nabla_{\boldsymbol{\xi}}\mathcal{H}, \quad (2.18)$$

układ równań różniczkowych zwyczajnych rzędu  $2M$  ze zmienną niezależną  $t$  i zmiennymi zależnymi  $\boldsymbol{\xi}$ . Jeśli wydzielić w wektorze stanu bloki zawierające współrzędne uogólnione  $\mathbf{q}$  i pędy uogólnione  $\mathbf{Q}$ , to stwierdzamy, że układ (2.18)

$$\begin{pmatrix} \dot{\mathbf{q}} \\ \dot{\mathbf{Q}} \end{pmatrix} = \begin{pmatrix} \mathbf{0} & \mathbf{E}_M \\ -\mathbf{E}_M & \mathbf{0} \end{pmatrix} \begin{pmatrix} \nabla_{\mathbf{q}}\mathcal{H} \\ \nabla_{\mathbf{Q}}\mathcal{H} \end{pmatrix},$$

można rozbić na

$$\dot{\mathbf{q}} = \nabla_{\mathbf{Q}}\mathcal{H}, \quad \dot{\mathbf{Q}} = -\nabla_{\mathbf{q}}\mathcal{H}, \quad (2.19)$$

albo, wypisując osobno pary dla każdego stopnia swobody  $i$

$$\frac{dq_i}{dt} = \frac{\partial\mathcal{H}}{\partial Q_i}, \quad \frac{dQ_i}{dt} = -\frac{\partial\mathcal{H}}{\partial q_i}, \quad i = 1, \dots, M. \quad (2.20)$$

Te eleganckie w swej symetrii i prostocie równania (2.18), (2.19) lub (2.20) nazywamy **równaniami kanonicznymi Hamiltona**. Dla zadanej funkcji  $\mathcal{H}$ , generującej prawe strony, ich rozwiązaniem są współrzędne i pędy uogólnione  $\mathbf{q}(t)$  i  $\mathbf{Q}(t)$  jako funkcje zmiennej niezależnej  $t$ .

## WYKŁAD 9

Wróćmy teraz do równania (2.16). Jakkolwiek nie wchodzi ono w skład równań kanonicznych, zawiera bardzo istotną konsekwencję postaci tych równań. Równanie (2.16) zawiera iloczyn skalarny, który wchodzi w skład różniczki zupełnej  $d\mathcal{H}$ . Rzeczywiście,

$$d\mathcal{H} = (D_{\mathbf{w}}\mathcal{H})d\mathbf{w} = (D_t\mathcal{H})dt + (D_{\boldsymbol{\xi}}\mathcal{H}) d\boldsymbol{\xi}.$$

A więc równanie (2.16) oznacza, że

$$d\mathcal{H} = (D_t\mathcal{H})dt,$$

i dzieląc przez  $dt$  otrzymujemy z niego wniosek

$$\frac{d\mathcal{H}}{dt} = \frac{\partial\mathcal{H}}{\partial t}. \quad (2.21)$$

W ten sposób udowodniliśmy następujące ważne twierdzenie:

**TWIERDZENIE 3** Jeżeli hamiltonian  $\mathcal{H}$  układu równań kanonicznych nie zależy od zmiennej niezależnej  $t$  w sposób jawny, to jest on całką pierwszą tego układu, czyli

$$\mathcal{H}(\mathbf{q}, \mathbf{Q}) = \text{const.} \quad (2.22)$$

dla każdego rozwiązania  $\mathbf{q}(t), \mathbf{Q}(t)$  równań kanonicznych.

Z samej postaci równań kanonicznych wynika kolejna ważna własność.

**TWIERDZENIE 4** Jeżeli hamiltonian  $\mathcal{H}$  pewnego układu nie zależy od jednej ze zmiennych kanonicznych, to sprzężona z nią zmienna jest stałą ruchu:

$$\begin{aligned} \frac{\partial\mathcal{H}}{\partial q_i} = 0 &\Rightarrow Q_i = \text{const.} \\ \frac{\partial\mathcal{H}}{\partial Q_i} = 0 &\Rightarrow q_i = \text{const.} \end{aligned}$$

Widać to wprost ze struktury równań (2.20).

Zmienna, która nie występuje w funkcji Hamiltona nosi nazwę zmiennej cyklicznej.

Podsumujmy: forma Pfaffa (2.15) generuje równania kanoniczne (2.18), które posiadają własności opisane twierdzeniami 3 i 4. Jeśli jednak przywołamy twierdzenie 2, to możemy podsumowanie nieco uogólnić:

Każdy pfaffian postaci

$$\Phi' = c \left( \mathcal{H} dt + \frac{[\boldsymbol{\xi} | d\boldsymbol{\xi}]}{2} + dS \right), \quad (2.23)$$

gdzie  $c$  jest dowolną stałą, a  $S = S(\boldsymbol{\xi}, t)$  dowolną funkcją różniczkowalną, generuje równania Pfaffa w postaci równań kanonicznych Hamiltona (2.18).

Z tego faktu można skorzystać, żeby pozbyć się połowy różniczek z pfaffianu. W większości prac przyjmuje się jako podstawowy pfaffian generujący równania kanoniczne formę, która powstaje z (2.23) dla  $c = -1$  i  $S = -\mathbf{q} \cdot \mathbf{Q}/2$ , co prowadzi do

$$\Phi' = -\mathcal{H} dt + \mathbf{Q} \cdot d\mathbf{q}. \quad (2.24)$$

Popularność tej formy Pfaffa bierze się z powiązań z optyką oraz faktem, że po wyrażeniu  $\mathbf{Q}$  przy pomocy  $\dot{\mathbf{q}}$ , jest ona równa funkcji Lagrange'a pomnożonej przez  $dt$ .

Można jednak wyeliminować z formy (2.15) różniczki  $d\mathbf{q}$ , podstawiając w (2.23)  $c = 1$  i  $S = \mathbf{q} \cdot \mathbf{Q}/2$ , dzięki czemu otrzymamy

$$\Phi' = \mathcal{H} dt + \mathbf{q} \cdot d\mathbf{Q}. \quad (2.25)$$

Formy Pfaffa w postaci (2.24) lub (2.25) (zawierające tylko połowę różniczek) nazywać będziemy **jednoformami kanonicznymi**.

## 2.4 Kanoniczne nawiasy Poissona

Niech  $\boldsymbol{\xi}$  oznacza  $2M$ -wymiarowy wektor stanu  $\text{col}(\mathbf{q}, \mathbf{Q})$  w przestrzeni fazowej, zaś  $F$  i  $G$  dowolne funkcje skalarne zmiennych  $\boldsymbol{\xi}$ . Załóżmy także, że ewolucja wektora stanu  $\boldsymbol{\xi}$  zadana jest funkcją Hamiltona  $\mathcal{H}$ .

Stawiamy teraz pytanie, jak wygląda pochodna dowolnej funkcji  $F(\boldsymbol{\xi}, t)$  względem zmiennej niezależnej  $t$ ? Korzystając z zasady różniczkowania funkcji złożonej uzyskujemy prosty wzór

$$\dot{F}(\boldsymbol{\xi}, t) = \nabla F \cdot \dot{\boldsymbol{\xi}} + \frac{\partial F}{\partial t} = D_{\boldsymbol{\xi}} F \dot{\boldsymbol{\xi}} + \frac{\partial F}{\partial t},$$

gdzie  $\nabla$  oznacza  $\nabla_{\boldsymbol{\xi}}$ . Sięgając do równań Hamiltona zastępujemy  $\dot{\boldsymbol{\xi}}$  przez prawą stronę (2.18) i otrzymujemy

$$\dot{F} = (\nabla F)^T \mathbf{J} \nabla \mathcal{H} + \frac{\partial F}{\partial t} = [\nabla F | \nabla \mathcal{H}] + \frac{\partial F}{\partial t}. \quad (2.26)$$

W powyższym wzorze pojawił się ważny obiekt matematyczny zwany **kanonicznym nawiasem Poissona**.

Kanonicznym nawiasem Poissona dwóch funkcji  $F$  i  $G$  nazywamy operator różniczkowy  $\{F, G\}$  zdefiniowany jako

$$\{F, G\} = (\nabla F)^T \mathbf{J} \nabla G = [\nabla F | \nabla G] = \sum_{i=1}^M \left( \frac{\partial F}{\partial q_i} \frac{\partial G}{\partial Q_i} - \frac{\partial F}{\partial Q_i} \frac{\partial G}{\partial q_i} \right). \quad (2.27)$$

Tam, gdzie będzie to ważne, dodawać będziemy do nawiasu Poissona indeks informujący o zmiennych kanonicznych względem których liczone są pochodne cząstkowe. Na przykład, we wzorze (2.27) mamy  $\{F, G\}_{\boldsymbol{\xi}}$ .

Wprowadzając nawias Poissona możemy zapisać równania kanoniczne (2.18) w jeszcze prostszej postaci

$$\dot{\boldsymbol{\xi}} = \{\boldsymbol{\xi}, \mathcal{H}\}, \quad (2.28)$$

a dla dowolnej funkcji  $F(\boldsymbol{\xi}, t)$ , w świetle równania (2.26),

$$\frac{dF}{dt} = \{F, \mathcal{H}\} + \frac{\partial F}{\partial t}. \quad (2.29)$$

Nawias Poissona posiada zarówno właściwości typowe dla każdego liniowego operatora różniczkowego (liniowość i działanie na iloczyn) jak i własności zbliżone do iloczynu wektorowego (antysymetria, tożsamość Jacobiego).

1. Antysymetryczność:

$$\{F, G\} = -\{G, F\}. \quad (2.30)$$

Oczywista konsekwencją antysymetryczności jest

$$\{F, F\} = 0.$$

2. Linowość. Dla dowolnych funkcji  $F, G, H$  i stałych  $\alpha, \beta$  zachodzi,

$$\{\alpha F + \beta G, H\} = \alpha \{F, H\} + \beta \{G, H\}. \quad (2.31)$$

W świetle własności (2.30) oznacza to w istocie dwuliniowość nawiasu Poissona (liniowość ze względu na oba argumenty).

3. Działanie na iloczyn funkcji:

$$\{FG, H\} = F\{G, H\} + G\{F, H\}. \quad (2.32)$$

Także i ta własność dotyczy również drugiego argumentu.

4. Tożsamość Jacobiego:

$$\{\{F, G\}, H\} + \{\{G, H\}, F\} + \{\{H, F\}, G\} = 0. \quad (2.33)$$

Dla porównania, dowolne trzy wektory  $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^3$  spełniają

$$(\mathbf{a} \times \mathbf{b}) \times \mathbf{c} + (\mathbf{b} \times \mathbf{c}) \times \mathbf{a} + (\mathbf{c} \times \mathbf{a}) \times \mathbf{b} = \mathbf{0}.$$

Właściwie należałoby stwierdzić, że każdy operator spełniający warunki 1-4 nazywamy nawiasem Poissona, natomiast szczególny przypadek (2.27) to kanoniczny nawias Poissona.

Warto też wiedzieć, że dla funkcji złożonej  $F(G(\boldsymbol{\xi}))$  mamy

$$\{F(G(\boldsymbol{\xi})), H(\boldsymbol{\xi})\} = F'(G) \{G, H\}, \quad (2.34)$$

a więc obowiązuje standardowa reguła różniczkowania funkcji złożonej.

## 2.5 Równania kanoniczne dla układu mechanicznego

### 2.5.1 Układ $K$ ciał

Równania kanoniczne Hamiltona pojawiają się w różnych dziedzinach matematyki powiązanych z teorią równań różniczkowych zwyczajnych i cząstkowych. Dla nas najważniejsze jest to, że mogą one opisywać ruch układu mechanicznego z czasem  $t$  jako zmienną niezależną. Pojawia się jednak istotne ograniczenie: formalizm kanoniczny można stosować tylko w układach, gdzie wszystkie działające siły są potencjalne, a więc tam, gdzie II zasada dynamiki ma postać

$$m_i \ddot{\mathbf{r}}_i = \mathbf{F}_i = -\nabla_{\mathbf{r}_i} V,$$

Przez  $V = V(\mathbf{r}_1, \dots, \mathbf{r}_K, t)$  oznaczyliśmy potencjał – skalarną funkcję, która generuje wszystkie siły w układzie  $K$  ciał. Używając wektora  $\mathbf{r} \in \mathbb{R}^{3K}$  w przestrzeni konfiguracyjnej

$$\mathbf{r} = \text{col}(\mathbf{r}_1, \dots, \mathbf{r}_K)^T,$$

mówimy, że  $V(\mathbf{r}, t)$  jest potencjałem zagadnienia, jeżeli wektor sił w przestrzeni konfiguracyjnej jest dany gradientem

$$\mathbf{F} = \text{col}(\mathbf{F}_1, \dots, \mathbf{F}_K)^T = -\nabla_{\mathbf{r}} V(\mathbf{r}, t). \quad (2.35)$$

Jeśli mamy do czynienia z zagadnieniem ruchu pod wpływem sił potencjalnych, to newtonowskie równania ruchu (1.2) w układzie inercjalnym można wyrazić w postaci kanonicznej poprzez prostą identyfikację:



- współrzędne uogólnione to wektor położeń w przestrzeni konfiguracyjnej  $\mathbf{q} = \mathbf{r}$ ,

$$r_1 = x_1, \quad r_2 = y_1, \quad r_3 = z_1, \quad r_4 = x_2, \quad r_5 = y_2, \quad r_6 = z_2, \quad \text{itd.}$$

- pędy uogólnione to pędy newtonowskie  $\mathbf{Q} = \mathbf{R}$  (iloczyn masy i prędkości)

$$R_1 = m_1 \dot{x}_1, \quad R_2 = m_1 \dot{y}_1, \quad R_3 = m_1 \dot{z}_1, \quad R_4 = m_2 \dot{x}_2, \quad \text{itd.}$$

czyli

$$\mathbf{R}_1 = m_1 \dot{\mathbf{r}}_1, \quad \mathbf{R}_2 = m_2 \dot{\mathbf{r}}_2, \quad \text{itd.}$$

(Nie należy mylić wektora pędu  $j$ -ego ciała ( $\mathbf{R}_j$ ) ze składową  $R_j$  pełnego wektora pędu  $\mathbf{R}$ .)

- funkcja Hamiltona równa jest energii całkowitej układu, czyli sumie energii kinetycznej  $T$  i potencjalnej  $V$  (terminu potencjał i energia potencjalna używać będziemy jako synonimów)

$$\mathcal{H}(\mathbf{r}, \mathbf{R}, t) = T(\mathbf{R}) + V(\mathbf{r}, t). \quad (2.36)$$

Dla układu  $K$  ciał energia kinetyczna jest sumą

$$T = \sum_{j=1}^K \frac{m_j \dot{\mathbf{r}}_j \cdot \dot{\mathbf{r}}_j}{2} = \sum_{j=1}^K \frac{1}{2m_j} \mathbf{R}_j \cdot \mathbf{R}_j. \quad (2.37)$$

Łatwo można sprawdzić, że przy tych założeniach forma Pfaffa

$$\Phi = \mathcal{H} dt + \frac{1}{2} \sum_{j=1}^K (\mathbf{r}_j \cdot d\mathbf{R}_j - \mathbf{R}_j \cdot d\mathbf{r}_j),$$

lub

$$\Phi' = -\mathcal{H} dt + \sum_{j=1}^K \mathbf{R}_j \cdot d\mathbf{r}_j,$$

generuje równania kanoniczne (2.19) w pełni równoważne układowi (1.2).

### 2.5.2 Zagadnienie dwóch ciał

Używając kartezjańskich zmiennych kanonicznych  $\mathbf{q} \in \mathbb{R}^6$  i  $\mathbf{Q} \in \mathbb{R}^6$  (bo  $K = 2$ ) w układzie inercyjnym

$$\mathbf{q} = \begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \end{pmatrix}, \quad \mathbf{Q} = \begin{pmatrix} \mathbf{R}_1 \\ \mathbf{R}_2 \end{pmatrix},$$

oraz funkcji Hamiltona

$$\mathcal{H} = \frac{R_1^2}{2m_1} + \frac{R_2^2}{2m_2} - \frac{k^2 m_1 m_2}{\|\mathbf{r}_2 - \mathbf{r}_1\|}, \quad (2.38)$$

gdzie  $k$  oznacza stałą Gaussa, otrzymujemy kanoniczne równania ruchu

$$\dot{\mathbf{q}} = \nabla_{\mathbf{Q}} \mathcal{H}, \quad \dot{\mathbf{Q}} = -\nabla_{\mathbf{q}} \mathcal{H},$$

czyli

$$\dot{\mathbf{r}}_1 = \frac{\mathbf{R}_1}{m_1}, \quad \dot{\mathbf{r}}_2 = \frac{\mathbf{R}_2}{m_2}, \quad (2.39)$$

$$\dot{\mathbf{R}}_1 = \frac{k^2 m_1 m_2}{\|\mathbf{r}_2 - \mathbf{r}_1\|^3} (\mathbf{r}_2 - \mathbf{r}_1), \quad \dot{\mathbf{R}}_2 = -\frac{k^2 m_1 m_2}{\|\mathbf{r}_2 - \mathbf{r}_1\|^3} (\mathbf{r}_2 - \mathbf{r}_1), \quad (2.40)$$

które są w pełni równoważne układowi znanemu z *Wstępu do mechaniki nieba*.

### 2.5.3 Ruch pojedynczej cząstki w polu sił

W szczególnym przypadku, gdy analizujemy ruch jednego ciała, możemy opuścić indeks numerujący ciała i dla zmiennych kanonicznych  $\mathbf{r}, \mathbf{R} \in \mathbb{R}^3$  mamy formę Pfaffa

$$\Phi = - \left( \frac{\mathbf{R} \cdot \mathbf{R}}{2m} + V(\mathbf{r}, t) \right) dt + \mathbf{R} \cdot d\mathbf{r},$$

z równaniami kanonicznymi

$$\dot{\mathbf{r}} = \frac{\mathbf{R}}{m}, \quad \dot{\mathbf{R}} = -\nabla_{\mathbf{r}} V(\mathbf{r}, t).$$

W tej sytuacji często korzystamy z Twierdzenia 2, w myśl którego mnożąc  $\Phi$  przez dowolną stałą otrzymamy równoważny układ równań. Wybierając  $c = m^{-1}$ , możemy zamiast  $\Phi$  użyć  $\Phi' = \Phi/m$ , czyli

$$\begin{aligned} \Phi' &= -\mathcal{H}'(\mathbf{r}, \mathbf{R}', t) dt + \mathbf{R}' \cdot d\mathbf{r}, \\ \mathbf{R}' &= \frac{\mathbf{R}}{m}, \\ \mathcal{H}'(\mathbf{r}, \mathbf{R}', t) &= \frac{\mathcal{H}}{m} = \frac{\mathbf{R}' \cdot \mathbf{R}'}{2} + V'(\mathbf{r}, t), \end{aligned} \quad (2.41)$$

gdzie  $V' = V/m$  to potencjał na jednostkę masy. Z tej formy Pfaffa z wektorem stanu  $\xi' = \text{col}(\mathbf{r}, \mathbf{R}')$  otrzymujemy równania kanoniczne typu (1.4)

$$\dot{\mathbf{r}} = \nabla_{\mathbf{R}'} \mathcal{H}' = \mathbf{R}', \quad \dot{\mathbf{R}}' = -\nabla_{\mathbf{r}} \mathcal{H}' = -\nabla_{\mathbf{r}} V'(\mathbf{r}, t), \quad (2.42)$$

z których widać, że pęd uogólniony  $\mathbf{R}'$  może być kartezjańskim wektorem prędkości, gdy funkcja Hamiltona ma charakter energii na jednostkę masy. Jest to reguła skalowania, z której chętnie korzystamy badając ruch jednego ciała, ale gdy ciał jest więcej, traci ona zalety, gdyż tylko jedna z mas może zostać użyta jako czynnik skali.

Z tego newtonowsko-kartezjańskiego punktu wyjścia będziemy przechodzić do coraz to ogólniejszych zmiennych  $(\mathbf{q}, \mathbf{Q})$  i niekoniecznie inercjalnych układów współrzędnych stosując transformacje kanoniczne, co jest typowym postępowaniem w ramach tego formalizmu.

## Rozdział 3

# Transformacje kanoniczne

WYKŁAD 10

### 3.1 Podstawy

Jeśli w ramach formalizmu newtonowskiego chcemy zastąpić zmienne kartezjańskie  $\mathbf{r}$  jakimiś ich funkcjami  $\mathbf{r}'(\mathbf{r}, t)$ , to w drugiej zasadzie dynamiki  $\ddot{\mathbf{r}} = \mathbf{F}$  zmiane ulegnie wszystko: i postać sił, i lewe strony. Wpływ przekształcenia zmiennych na postać równań ruchu jest trudny do przewidzenia z góry. Czy równania będą prostsze, czy bardziej złożone? Tego nie wiemy, póki ich nie wyprowadzimy. Formalizm kanoniczny oferuje nam tu o wiele bardziej komfortową sytuację. Dopiero w jego ramach przekształcenia zmiennych stają się sprawnym narzędziem upraszczania równań ruchu.

#### 3.1.1 Definicja transformacji kanonicznej

Pojęcie transformacji kanonicznej odgrywa fundamentalną rolę w mechanice hamiltonowskiej. Wiąże się ono z następującym problemem:

Mamy dany układ mechaniczny zdefiniowany przez funkcję Hamiltona  $\mathcal{H}$ , wektor stanu (zmiennie)  $\boldsymbol{\xi}$  i kanoniczne równania ruchu. Chcemy do opisu ruchu tego układu użyć innych zmiennych  $\boldsymbol{\eta}$ , które można wyrazić przy pomocy  $\boldsymbol{\xi}$  jako wartości funkcji  $\mathbf{Y}$ . Dopuszczając jawną zależność przekształcenia od czasu, mamy  $\boldsymbol{\eta} = \mathbf{Y}(\boldsymbol{\xi}, t)$ .

Jakie warunki muszą spełnić funkcje  $\mathbf{Y}$ , aby równania ruchu w zmiennych  $\boldsymbol{\eta}$  były nadal kanoniczne?

Takie przekształcenie zmiennych (wektora stanu), które zachowuje kanoniczną postać równań ruchu nazywamy **transformacją kanoniczną**.

Z punktu widzenia równań Pfaffa, można powiedzieć, że transformacja jest kanoniczna, gdy wyrażając pffaffian

$$\Phi = \mathcal{H}dt + \frac{1}{2} [\boldsymbol{\xi} | d\boldsymbol{\xi}],$$

przy pomocy nowych zmiennych  $\boldsymbol{\eta}$ , otrzymamy formę

$$\Phi' = \mathcal{K}dt + \frac{1}{2} [\boldsymbol{\eta} | d\boldsymbol{\eta}],$$

jak zwykle – z dokładnością do mnożnika i różniczki zupełnej. Jest to definicja w pełni równoważna poprzedniej. Niestety, bez wejścia w teorię różniczkowych form zewnętrznych trudno wydobyć z niej ogólny warunek kanoniczności transformacji, choć w prostszych wypadkach (np. przekształcenia liniowe) będzie ona poręczna.

Z tych powodów, warunek dostateczny kanoniczności transformacji otrzymamy badając równania ruchu.

### 3.1.2 Warunek dostateczny kanoniczności transformacji

Niech  $\boldsymbol{\xi} = \text{col}(\mathbf{q}, \mathbf{Q})$  oznacza wektor stanu, który nazwiemy „starymi zmiennymi” kanonicznymi. Ewolucja starych zmiennych opisana jest równaniami kanonicznymi (2.18)

$$\frac{d\boldsymbol{\xi}}{dt} = \mathbf{J} \nabla_{\boldsymbol{\xi}} \mathcal{H},$$

z funkcją Hamiltona  $\mathcal{H}(\boldsymbol{\xi}, t)$ .

Wprowadzamy teraz nowy wektor stanu  $\boldsymbol{\eta} = \text{col}(\mathbf{p}, \mathbf{P})$  i potrafimy wyrazić go przy pomocy starych zmiennych, dopuszczając jawną zależność od czasu (na przykład, jeśli nowe zmienne odniesione są do układu nieinercjalnego)

$$\boldsymbol{\eta} = \mathbf{Y}(\boldsymbol{\xi}, t).$$

W powyższym wzorze staramy się odróżnić zmienne  $\boldsymbol{\eta}$  od funkcji  $\mathbf{Y}$ , które określają ich zależność od  $\boldsymbol{\xi}$  i jawną od czasu  $t$ . Przekształcenie powinno być odwracalne, to znaczy musi istnieć funkcja odwrotna  $\mathbf{Y}^{-1}$ , umożliwiająca powrót z nowych zmiennych do starych

$$\boldsymbol{\xi} = \mathbf{Y}^{-1}(\boldsymbol{\eta}, t).$$

Spytajmy teraz o pochodną  $\boldsymbol{\eta}$  względem czasu. Jeśli wprowadzimy macierz Jacobiego

$$\mathbf{U} = D_{\boldsymbol{\xi}}\mathbf{Y},$$

to

$$\frac{d\boldsymbol{\eta}}{dt} = \mathbf{U} \frac{d\boldsymbol{\xi}}{dt} + D_t\mathbf{Y}.$$

Ponieważ ewolucja  $\boldsymbol{\xi}$  opisana jest równaniami (2.18), mamy

$$\dot{\boldsymbol{\eta}} = \mathbf{U}\mathbf{J}\nabla_{\boldsymbol{\xi}}\mathcal{H} + D_t\mathbf{Y}.$$

W tym równaniu nadal występują stare zmienne i to w trzech miejscach: w operatorze gradientu, wewnątrz hamiltonianu  $\mathcal{H}$  i w funkcjach  $\mathbf{Y}$ . Z pierwszymi dwoma radzimy sobie prosto:

- zamiana zmiennych w gradiencie sprowadza się do

$$\nabla_{\boldsymbol{\xi}} = \mathbf{U}^T \nabla_{\boldsymbol{\eta}}. \quad (3.1)$$

- Funkcję Hamiltona  $\mathcal{H}(\boldsymbol{\xi}, t)$  możemy wyrazić przy pomocy nowych zmiennych podstawiając wzory transformacji odwrotnej

$$\mathcal{H}(\boldsymbol{\xi}, t) = \mathcal{H}(\mathbf{Y}^{-1}(\boldsymbol{\eta}, t), t) = \mathcal{H}'(\boldsymbol{\eta}, t). \quad (3.2)$$

Daje to nam gwarancję, że opisujemy w nowych zmiennych ruch tego samego układu.

W ten sposób sprowadziliśmy równania dla  $\dot{\boldsymbol{\eta}}$  do postaci

$$\dot{\boldsymbol{\eta}} = \mathbf{U}\mathbf{J}\mathbf{U}^T \nabla_{\boldsymbol{\eta}}\mathcal{H}' + D_t\mathbf{Y}.$$

Przywołując równanie (2.9) zauważamy, że jeśli macierz Jacobiego  $\mathbf{U}$  jest macierzą symplektyczną to równania ruchu w nowych zmiennych przybierają postać

$$\dot{\boldsymbol{\eta}} = \mathbf{J}\nabla_{\boldsymbol{\eta}}\mathcal{H}' + D_t\mathbf{Y}. \quad (3.3)$$

Gdyby nie jawna zależność transformacji od czasu, to z  $D_t\mathbf{Y} = \mathbf{0}$  mielibyśmy równania kanoniczne w nowych zmiennych z funkcją Hamiltona  $\mathcal{H}'(\boldsymbol{\eta}, t)$ .

Żeby usunąć wyraz  $D_t\mathbf{Y}$ , który nadal pozostaje funkcją starych zmiennych, wprowadzamy do funkcji Hamiltona tak zwaną **resztę transformacji**  $\mathcal{R}(\boldsymbol{\eta}, t)$ . Przyjmujemy, że nową funkcją Hamiltona będzie

$$\mathcal{K}(\boldsymbol{\eta}, t) = \mathcal{H}'(\boldsymbol{\eta}, t) + \mathcal{R}(\boldsymbol{\eta}, t), \quad (3.4)$$

i podstawiamy  $\mathcal{H}' = \mathcal{K} - \mathcal{R}$  do równania (3.3), otrzymując

$$\dot{\boldsymbol{\eta}} = \mathbf{J} \nabla_{\boldsymbol{\eta}} \mathcal{K} - [\mathbf{J} \nabla_{\boldsymbol{\eta}} \mathcal{R} - D_t \mathbf{Y}].$$

Wyraz w nawiasie kwadratowym można sprowadzić do zera znajdując odpowiednią resztę  $\mathcal{R}$  jako rozwiązanie równania różniczkowego cząstkowego

$$\nabla_{\boldsymbol{\eta}} \mathcal{R} = -\mathbf{J} \mathbf{Y}'(\boldsymbol{\eta}, t). \quad (3.5)$$

Przez  $\mathbf{Y}'(\boldsymbol{\eta}, t)$  oznaczyliśmy wektor który powstaje w następujący sposób: znajdujemy pochodną cząstkową  $\mathbf{Y}_t = D_t \mathbf{Y}(\boldsymbol{\xi}, t)$ , a następnie zmienne  $\boldsymbol{\xi}$  w  $\mathbf{Y}_t$  wyrażamy przy pomocy nowych zmiennych  $\boldsymbol{\eta}$ , czyli

$$\mathbf{Y}'(\boldsymbol{\eta}, t) = [D_t \mathbf{Y}(\boldsymbol{\xi}, t)]_{\boldsymbol{\xi}=\mathbf{Y}^{-1}(\boldsymbol{\eta}, t)}.$$

Nie będziemy rozpatrywać bliżej sposobu znajdowania reszty, gdyż dalej poznamy szczególne rodzaje transformacji kanonicznych, dla których istnieją proste reguły otrzymywania  $\mathcal{R}$ .

Przedstawione wyżej wyprowadzenie prowadzi do sformułowania następującego warunku kanoniczności transformacji:

Transformacja  $(\boldsymbol{\xi}, \mathcal{H}) \rightarrow (\boldsymbol{\eta}, \mathcal{K})$  zachowuje kanoniczną postać równań ruchu gdy:

- macierz Jacobiego  $\mathbf{U} = D_{\boldsymbol{\xi}} \boldsymbol{\eta}$  (poprawniej:  $D_{\boldsymbol{\xi}} \mathbf{Y}(\boldsymbol{\xi}, t)$ ) jest symplektyczna,
- hamiltonian  $\mathcal{K}$  powstaje przez podstawienie do  $\mathcal{H}$  związków między zmiennymi i – dla transformacji jawnie zależnych od czasu – dodanie reszty transformacji.

Przedstawione kryterium warunek ma charakter warunku dostatecznego, ale nie koniecznego. Istnieje możliwość otrzymania równań kanonicznych w nowych zmiennych przez dodatkowe zabiegi na funkcji Hamiltona bez spełnienia warunku symplektyczności  $\mathbf{U}$ . Najprostszym przykładem jest skalowanie zmiennych z rozdziału 2.5.3 (dodatkowym zabiegiem było podzielenie przez masę), które zachowuje postać kanoniczną równań ruchu, choć jego macierz Jacobiego nie jest symplektyczna.

Symplektyczność macierzy Jacobiego gwarantuje, że transformacje kanoniczne tworzą grupę, w której działaniem (nieprzemienne) jest złożenie transformacji. W szczególności, transformacja odwrotna do kanonicznej jest kanoniczna.

### 3.1.3 Nawiasy Poissona a transformacja kanoniczna

Warunek symplektyczności macierzy Jacobiego  $\mathbf{U}$  łatwo można powiązać z nawiasami Poissona. Załóżmy, że mamy zmienne  $\boldsymbol{\xi}$  i  $\boldsymbol{\eta} = \mathbf{Y}(\boldsymbol{\xi})$  powiązane transformacją kanoniczną. W rozważaniach tego rozdziału jawna zależność od czasu lub jej brak są całkowicie bez znaczenia, więc dla zwięzłości pomijamy  $t$ .

Rozpatrzmy nawias Poissona dwóch funkcji nowych zmiennych  $F(\boldsymbol{\eta})$ ,  $G(\boldsymbol{\eta})$ . Zgodnie z definicją (2.27), nawias Poissona tych funkcji to

$$\{F, G\}_{\boldsymbol{\eta}} = [\nabla_{\boldsymbol{\eta}} F \mid \nabla_{\boldsymbol{\eta}} G].$$

Wyraźmy teraz  $F$  i  $G$  przy pomocy starych zmiennych  $\boldsymbol{\xi}$ , otrzymując

$$F'(\boldsymbol{\xi}) = (F \circ \mathbf{Y})(\boldsymbol{\xi}) = F(\mathbf{Y}(\boldsymbol{\xi})),$$

$$G'(\boldsymbol{\xi}) = (G \circ \mathbf{Y})(\boldsymbol{\xi}) = G(\mathbf{Y}(\boldsymbol{\xi})).$$

Zgodnie z regułą transformacji gradientu (3.1), pomnożoną stronami przez macierz odwrotną  $(\mathbf{U}^T)^{-1}$ , mamy

$$\{F, G\}_{\boldsymbol{\eta}} = [(\mathbf{U}^T)^{-1} \nabla_{\boldsymbol{\xi}} F' \mid (\mathbf{U}^T)^{-1} \nabla_{\boldsymbol{\xi}} G'].$$

Ale jeśli  $U$  jest macierzą symplektyczną, to zarówno  $U^T$ , jak i macierz odwrotna do niej są symplektyczne, więc nie zmieniają wartości formy symplektycznej w myśl (2.5), zatem

$$\{F, G\}_{\boldsymbol{\eta}} = [\nabla_{\boldsymbol{\xi}} F' \mid \nabla_{\boldsymbol{\xi}} G'] = \{F', G'\}_{\boldsymbol{\xi}}.$$

I tak doszliśmy do zasady niezmienniczości nawiasów Poissona przy transformacji kanonicznej  $\boldsymbol{\eta} = \mathbf{Y}(\boldsymbol{\xi})$ :

$$\{F(\boldsymbol{\eta}), G(\boldsymbol{\eta})\}_{\boldsymbol{\eta}} = \{(F \circ \mathbf{Y})(\boldsymbol{\xi}), (G \circ \mathbf{Y})(\boldsymbol{\xi})\}_{\boldsymbol{\xi}}. \quad (3.6)$$

Można też zapisać ten warunek w skróconej postaci

$$\{F, G\}_{\boldsymbol{\eta}} = \{F, G\}_{\boldsymbol{\xi}}, \quad (3.7)$$

pamiętając, że pochodne cząstkowe  $F$  i  $G$  muszą być liczone albo po podstawieniu równań transformacji, albo przy użyciu wzoru na pochodną funkcji złożonej.



Łatwo sprawdzić, że dla kanonicznych zmiennych  $\boldsymbol{\eta} = \text{col}(\mathbf{p}, \mathbf{P})$

$$\{p_i, P_j\}_{\boldsymbol{\eta}} = \delta_{ij}, \quad \{p_i, p_j\}_{\boldsymbol{\eta}} = \{P_i, P_j\}_{\boldsymbol{\eta}} = 0, \quad (3.8)$$

gdzie  $\delta_{ij}$  oznacza deltę Kroneckera. Podobnie dla  $\boldsymbol{\xi} = \text{col}(\mathbf{q}, \mathbf{Q})$

$$\{q_i, Q_j\}_{\boldsymbol{\xi}} = \delta_{ij},$$

a pozostałe nawiasy są zerowe. W zwartej postaci formułujemy tę własność mówiąc, że macierz nawiasów Poissona zmiennych kanonicznych jest standardową macierzą symplektyczną

$$\mathbf{M} = \{\boldsymbol{\eta}, \boldsymbol{\eta}\}_{\boldsymbol{\eta}} = \mathbf{J}, \quad (3.9)$$

gdzie  $M_{ij} = \{\eta_i, \eta_j\}_{\boldsymbol{\eta}}$ .

Niezmienniczość nawiasów Poissona pozwala więc zastąpić warunek symplektyczności macierzy Jacobiego transformacji  $\boldsymbol{\eta} = \mathbf{Y}(\boldsymbol{\xi}, t)$ , przez następujące kryterium:

jeśli

$$\{\mathbf{Y}(\boldsymbol{\xi}, t), \mathbf{Y}(\boldsymbol{\xi}, t)\}_{\boldsymbol{\xi}} = \mathbf{J}, \quad (3.10)$$

to transformacja jest kanoniczna.

### Przykład: Transformacja Poincarégo dla oscylatora harmonicznego

Rozpatrzmy Hamiltonian oscylatora harmonicznego

$$\mathcal{H} = \frac{1}{2} (X^2 + \omega^2 x^2). \quad (3.11)$$

Równania ruchu dla zmiennych  $\boldsymbol{\xi} = (x, X)^T$  mają znaną postać

$$\dot{x} = \{x, \mathcal{H}\} = X, \quad \dot{X} = \{X, \mathcal{H}\} = -\omega^2 x.$$

Można jednak wprowadzić nowe zmienne kanoniczne  $\boldsymbol{\eta} = (\ell, L)^T$ , dla których nowy hamiltonian  $\mathcal{K}$  będzie miał prostszą postać. Są to tzw. zmienne Poincaré, związane z  $x, X$  poprzez

$$x = \sqrt{\frac{2L}{\omega}} \sin \ell, \quad X = \sqrt{2L\omega} \cos \ell. \quad (3.12)$$

Sprawdźmy kanoniczność tych zmiennych badając nawiasy Poissona:

$$\{x, X\}_{\boldsymbol{\eta}} = \frac{\partial x}{\partial \ell} \frac{\partial X}{\partial L} - \frac{\partial x}{\partial L} \frac{\partial X}{\partial \ell} = 1,$$

zaś  $\{x, x\} = \{X, X\} = 0$  z definicji nawiasów (antysymetryczność!). Skoro kryterium (3.10) zostało spełnione, możemy uznać nowe zmienne za kanoniczne i znaleźć nowy hamiltonian. Transformacja nie zależy jawnie od czasu, więc nie wprowadzamy reszty transformacji i

$$\mathcal{K} = \mathcal{H}(x(\ell, L), X(\ell, L)) = \omega L. \quad (3.13)$$

Zauważmy, że nowy hamiltonian nie zależy od zmiennej  $\ell$ ; dla podkreślenia tego faktu piszemy zwykle

$$\mathcal{K} = \mathcal{K}(-, L).$$

Nowe równania ruchu

$$\dot{\ell} = \omega, \quad \dot{L} = 0, \quad (3.14)$$

są na tyle proste, że można je rozwiązać bez trudu

$$L = L_0 = \text{const}, \quad \ell = \omega(t - t_0) + \ell_0. \quad (3.15)$$

Pojawiły się dwie stałe dowolne,  $L_0$  i  $\ell_0$ , które wyznaczamy z warunków początkowych.

Przykład ten dobrze ilustruje rolę transformacji kanonicznych w upraszczaniu równań ruchu.

## 3.2 Transformacje Mathieu

### 3.2.1 Definicja

Najprostszy rodzaj transformacji kanonicznej to transformacje Mathieu. W najbardziej ogólnej postaci,

**transformacja Mathieu** to takie kanoniczne przekształcenie zmiennych, które zachowuje wartość i postać kanonicznej formy (2.24) lub (2.25).

Oczywiście, warunek zachowania wartości formy wyklucza dodawanie różniczki zupełnej oraz stosowanie mnożnika innego niż 1, mimo że nie wpływa to na równania kanoniczne ruchu. Ale takie ograniczenie znacząco upraszcza praktyczne wykonanie transformacji Mathieu.

Najczęstszy wariant transformacji Mathieu polega na zadaniu nowych współrzędnych jako funkcji współrzędnych starych i czasu, a potem takim doborze nowych pędów, żeby zagwarantować kanoniczność (wariant ten znany jest jako *rozszerzenie kanoniczne transformacji punktowej*). Ze względu na symetrię współrzędnych i pędów uogólnionych, można także prowadzić transformacje Mathieu zadając nowe pędy jako funkcje starych (i czasu) a potem dobierać nowe współrzędne.

Przyjmijmy za punkt wyjścia zmienne kanoniczne  $\xi = \text{col}(\mathbf{q}, \mathbf{Q})$  dla układu z funkcją Hamiltona  $\mathcal{H}$ . W wariacie podstawowym wykonujemy najpierw transformację punktową (nazwa myląca – nie ma nic wspólnego z pojęciem grupy punktowej transformacji w algebrze), która polega na wprowadzeniu nowych współrzędnych  $\mathbf{p}$ , zależnych jedynie od starych współrzędnych  $\mathbf{q}$  i jawnie od czasu

$$\mathbf{p} = \mathbf{p}(\mathbf{q}, t). \quad (3.16)$$

Zauważmy, że jest to wąska klasa przekształceń, zabraniająca zależności  $\mathbf{p}$  od starych pędów  $\mathbf{Q}$  (przedstawiona wyżej transformacja Poincaré nie jest transformacją punktową). Rozszerzenie kanoniczne transformacji punktowej (3.16) to znalezienie pędów  $\mathbf{P}$  sprzężonych kanonicznie z nowymi współrzędnymi oraz nowego hamiltonianu  $\mathcal{K}$ .

Rozszerzenie kanoniczne wykonujemy w oparciu o warunek zachowania wartości i postaci formy różniczkowej Pfaffa (2.24)

$$-\mathcal{H} dt + \mathbf{Q}^T d\mathbf{q} = -\mathcal{K} dt + \mathbf{P}^T d\mathbf{p}. \quad (3.17)$$

Zauważmy od razu, że jeśli transformacja (3.16) nie zależy jawnie od czasu, to nie pojawi się różniczka  $dt$  wynikająca z  $d\mathbf{p}$  i będziemy mieli  $\mathcal{H} = \mathcal{K}$  a więc transformację bez reszty  $\mathcal{R}$ .

Wariant alternatywny (rzadszy) to wybór nowych współrzędnych zachowujący wartość jednoformy (2.25), czyli

$$\mathcal{H} dt + \mathbf{q}^T d\mathbf{Q} = \mathcal{K} dt + \mathbf{p}^T d\mathbf{P}. \quad (3.18)$$

Zaletą transformacji Mathieu jest prostota ich realizacji oraz automatyczne generowanie reszty transformacji.

### 3.2.2 Liniowa transformacja kanoniczna

Założmy, że w układzie kanonicznym z funkcja Hamiltona  $\mathcal{H}(\mathbf{q}, \mathbf{Q}, t)$  wprowadzimy nowe współrzędne uogólnione  $\mathbf{p}$  poprzez liniową transformację punktową

$$\mathbf{p} = \mathbf{A}\mathbf{q}, \quad (3.19)$$

gdzie nieosobliwa macierz kwadratowa nie zależy od  $\mathbf{q}$  ani  $\mathbf{Q}$ , a jedynie od czasu. Wtedy

$$d\mathbf{p} = \mathbf{A}d\mathbf{q} + \dot{\mathbf{A}}\mathbf{q} dt.$$

A zatem, warunek transformacji Mathieu (3.17) przybiera postać

$$-\mathcal{H} dt + \mathbf{Q}^T d\mathbf{q} = -\mathcal{K} dt + \mathbf{P}^T \dot{\mathbf{A}}\mathbf{q} dt + \mathbf{P}^T \mathbf{A}d\mathbf{q}. \quad (3.20)$$

Przyrównując wyrazy mnożone przez  $d\mathbf{q}$  stwierdzamy, że

$$\mathbf{P}^T \mathbf{A} = \mathbf{Q}^T,$$

czyli, transponując obie strony,

$$\mathbf{A}^T \mathbf{P} = \mathbf{Q},$$

co prowadzi do definicji nowych pędów

$$\mathbf{P} = \mathbf{A}^{-T} \mathbf{Q}. \quad (3.21)$$

Przez  $\mathbf{A}^{-\text{T}}$  rozumiemy odwrotność macierzy transponowanej albo transpozycję macierzy odwrotnej, co na jedno wychodzi, gdyż

$$(\mathbf{A}^{-1})^{\text{T}} = (\mathbf{A}^{\text{T}})^{-1} = \mathbf{A}^{-\text{T}}.$$

Nowy hamiltonian  $\mathcal{K}$  otrzymamy przyrównując wyrazy (3.20) mnożone przez  $dt$

$$-\mathcal{H} = -\mathcal{K} + \mathbf{P}^{\text{T}} \dot{\mathbf{A}} \mathbf{q},$$

czyli

$$\mathcal{K} = \mathcal{H} + \mathbf{P}^{\text{T}} \dot{\mathbf{A}} \mathbf{q}.$$

Oczywiście, i stary hamiltonian  $\mathcal{H}$ , i reszta transformacji muszą być wyrażone przy pomocy nowych zmiennych, więc podstawiając rozwiązania (3.19) i (3.21) względem starych współrzędnych i pędów

$$\mathbf{q} = \mathbf{A}^{-1} \mathbf{p}, \quad \mathbf{Q} = \mathbf{A}^{\text{T}} \mathbf{P},$$

otrzymamy

$$\mathcal{K}(\mathbf{p}, \mathbf{P}, t) = \mathcal{H}(\mathbf{A}^{-1} \mathbf{p}, \mathbf{A}^{\text{T}} \mathbf{P}, t) + \mathbf{P}^{\text{T}} \dot{\mathbf{A}} \mathbf{A}^{-1} \mathbf{p}, \quad (3.22)$$

jako nową funkcję Hamiltona.

Możemy więc sformułować następujące twierdzenie o liniowej transformacji kanonicznej

**TWIERDZENIE 5:** Niech  $\boldsymbol{\xi}, \boldsymbol{\eta} \in \mathbb{R}^{2M}$  oznaczają dwa wektory stanu  $\boldsymbol{\xi} = \text{col}(\mathbf{q}, \mathbf{Q})$ ,  $\boldsymbol{\eta} = \text{col}(\mathbf{p}, \mathbf{P})$ , zaś  $\mathbf{A} = \mathbf{A}(t)$  nieosobliwą macierz kwadratową stopnia  $M$ . Transformacja liniowa

$$\boldsymbol{\eta} = \mathbf{S} \boldsymbol{\xi},$$

z macierzą

$$\mathbf{S} = \begin{pmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}^{-\text{T}} \end{pmatrix}, \quad (3.23)$$

jest transformacją kanoniczną a jej reszta wynosi

$$\mathcal{R} = \mathbf{P}^{\text{T}} \dot{\mathbf{A}} \mathbf{A}^{-1} \mathbf{p}. \quad (3.24)$$

### 3.2.3 Ruch w układzie jednostajnie obracających się osi

Rozpatrzmy często spotykaną w mechanice sytuację, gdy chcemy badać ruch w jednostajnie obracającym się układzie współrzędnych. Punktem wyjścia będą położenia i pędy (na jednostkę masy) w kartezjańskim układzie inercyjnym  $Ox_0y_0z_0$

$$\boldsymbol{\xi} = \text{col}(\mathbf{r}_0, \mathbf{R}_0),$$

gdzie,

$$\mathbf{r}_0 = (x_0, y_0, z_0)^T, \quad \mathbf{R}_0 = (X_0, Y_0, Z_0)^T.$$

Funkcja Hamiltona w tym układzie ma postać

$$\mathcal{H}(\mathbf{r}_0, \mathbf{R}_0, t) = \frac{1}{2} \mathbf{R}_0^T \mathbf{R}_0 + V(\mathbf{r}_0, t), \quad (3.25)$$

gdzie  $V$  oznacza dowolny potencjał (na jednostkę masy).

Wprowadźmy teraz drugi układ współrzędnych  $Oxyz$  o tym samym środku  $O$ , który obraca się jednostajnie wokół osi  $Oz_0 = Oz$ , z prędkością kątową  $\Omega$ . Chcemy, aby nowe zmienne położenia opisywały pozycję punktu materialnego w układzie rotującym, a więc oznaczając przez  $\mathbf{M}_3(\alpha)$  macierz obrotu pasywnego o kąt  $\alpha$  wokół osi  $Oz$  (pasywnego, gdyż nie obracamy wektora  $\mathbf{r}$ , lecz jego współrzędne ulegają zmianie z powodu obrotu układu współrzędnych), zadajemy transformację w postaci

$$\mathbf{r} = \mathbf{M}_3(\Omega t) \mathbf{r}_0, \quad (3.26)$$

czyli

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \cos(\Omega t) & \sin(\Omega t) & 0 \\ -\sin(\Omega t) & \cos(\Omega t) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix}.$$

Jak widać, związek (3.26) definiuje transformację punktową, gdyż nie pojawiają się w nim „stare pędy”  $\mathbf{R}_0$  a tylko współrzędne  $\mathbf{r}_0$  i czas  $t$ . Jest to w istocie przypadek omawiany w rozdziale 3.2.2 i można do niego zastosować Twierdzenie 5. Rolę macierz  $\mathbf{A}$  pełni tu  $\mathbf{M}_3(\Omega t)$ , więc według wzoru (3.23)

$$\mathbf{R} = \mathbf{M}_3^{-T} \mathbf{R}_0,$$

ale mamy do czynienia z macierzą ortogonalną, odwrotność równa jest transpozycji, co oznacza

$$\mathbf{R} = \mathbf{M}_3(\Omega t) \mathbf{R}_0. \quad (3.27)$$

Transformacja pędów okazuje się identyczna jak transformacja współrzędnych.

Transformacja jest jawnie zależna od czasu, więc nowy Hamiltonian dany jest jako

$$\mathcal{K}(\mathbf{r}, \mathbf{R}, t) = \mathcal{H}(\mathbf{M}_3(-\Omega t)\mathbf{r}, \mathbf{M}_3(-\Omega t)\mathbf{R}, t) + \mathcal{R}, \quad (3.28)$$

zaś reszta transformacji, zgodnie z równaniem (3.24), ma postać

$$\mathcal{R} = \mathbf{R}^T \dot{\mathbf{M}}_3(\Omega t) \mathbf{M}_3(-\Omega t) \mathbf{r}.$$

Ponieważ dla macierzy obrotu

$$\mathbf{M}_3(\Omega t) = \begin{pmatrix} \cos \Omega t & \sin \Omega t & 0 \\ -\sin \Omega t & \cos \Omega t & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

pochodna wynosi

$$\dot{\mathbf{M}}_3(\Omega t) = \Omega \begin{pmatrix} -\sin \Omega t & \cos \Omega t & 0 \\ -\cos \Omega t & -\sin \Omega t & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad (3.29)$$

to

$$\dot{\mathbf{M}}_3(\Omega t) \mathbf{M}_3(-\Omega t) = \Omega \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

a więc

$$\mathcal{R} = -\Omega (xY - yX).$$

Zauważmy, że wyrażenie w nawiasie to nic innego jak składowa z wektora momentu pędu (rzut momentu pędu na oś obrotu).

Związek (3.27) oznacza, że długości starego i nowego wektora pędu są jednakowe, a więc uwzględniając  $\mathbf{R}_0^T \mathbf{R}_0 = \mathbf{R}^T \mathbf{R}$ , sięgamy do (3.25) aby otrzymać końcową postać funkcji Hamiltona w obracającym się układzie współrzędnych

$$\mathcal{K} = \frac{1}{2} \mathbf{R}^T \mathbf{R} - \Omega (xY - yX) + V'(\mathbf{r}, t), \quad (3.30)$$

gdzie  $V'(\mathbf{r}, t) = V(\mathbf{M}_3(-\Omega t)\mathbf{r}, t)$ .

Podsumujmy krótko wyniki tego ważnego przykładu:

- Transformacja (3.26) jest jawnie zależna od czasu a więc  $\mathcal{K} \neq \mathcal{H}$ . Reszta transformacji ma postać  $\mathcal{R} = -\Omega (xY - yX)$ .

- Pędy  $\mathbf{R}$  w rotującym układzie współrzędnych nie są tożsame z prędkościami w tym układzie (oprócz  $Z = \dot{z}$ ), gdyż

$$\begin{aligned}\dot{x} &= \frac{\partial \mathcal{K}}{\partial X} = X + \Omega y, \\ \dot{y} &= \frac{\partial \mathcal{K}}{\partial Y} = Y - \Omega x, \\ \dot{z} &= \frac{\partial \mathcal{K}}{\partial Z} = Z.\end{aligned}\tag{3.31}$$

Cząstka obracająca się wraz z układem ma w nim zerową prędkość, ale nie zerowy pęd.

- Reszta transformacji powoduje pojawienie się „sił pozornych” oprócz tych, wynikających z gradientu potencjału:

$$\begin{aligned}\dot{X} &= -\frac{\partial \mathcal{K}}{\partial x} = \Omega Y - \frac{\partial V'}{\partial x}, \\ \dot{Y} &= -\frac{\partial \mathcal{K}}{\partial y} = -\Omega X - \frac{\partial V'}{\partial y}, \\ \dot{Z} &= -\frac{\partial \mathcal{K}}{\partial z} = -\frac{\partial V'}{\partial z}.\end{aligned}\tag{3.32}$$

- Nowe pędy są w istocie pędami z układu inercyjnego, których wektor jest jedynie obracany o kąt  $\Omega t$  (czyli rzutowany na chwilowe osie).



### 3.2.4 Zagadnienie względne dwóch ciał

Wychodząc od pełnego zagadnienia dwóch ciał z rozdziału 2.5.2, przejdziemy do kanonicznego sformułowania zagadnienia względnego, czyli ruchu masy  $m_2$  względem masy  $m_1$ . Zobaczymy, że i tu można zastosować transformację Mathieu oraz Twierdzenie 5.

Niech pierwotne zmienne kanoniczne  $\xi$  dane są przez

$$\xi = \begin{pmatrix} \mathbf{q} \\ \mathbf{Q} \end{pmatrix} = \begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \mathbf{R}_1 \\ \mathbf{R}_2 \end{pmatrix},$$

gdzie  $\mathbf{R}_i = m_i \dot{\mathbf{r}}_i$ , zaś funkcja Hamiltona  $\mathcal{H}$  ma postać (2.38). Poszukujemy teraz takich zmiennych kanonicznych  $\eta$ , w których zamiast  $\mathbf{r}_2$  pojawi się położenie względne  $\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1$ . Drugim wektorem położenia niech będzie wektor środka masy  $\mathbf{r}_0 = (m_1 \mathbf{r}_1 + m_2 \mathbf{r}_2)/(m_1 + m_2)$ . A zatem, nowe współrzędne uogólnione zadane są poprzez niezależną od czasu, liniową transformację punktową. Niech

$$m = \frac{m_1}{m_1 + m_2}, \quad 1 - m = \frac{m_2}{m_1 + m_2}. \quad (3.33)$$

Mamy wtedy

$$\mathbf{p} = \begin{pmatrix} \mathbf{r} \\ \mathbf{r}_0 \end{pmatrix} = \begin{pmatrix} -\mathbf{r}_1 + \mathbf{r}_2 \\ m\mathbf{r}_1 + (1 - m)\mathbf{r}_2 \end{pmatrix} = \mathbf{A} \mathbf{q}, \quad (3.34)$$

gdzie

$$\mathbf{A} = \begin{pmatrix} -\mathbf{E} & \mathbf{E} \\ m\mathbf{E} & (1 - m)\mathbf{E} \end{pmatrix},$$

zaś  $\mathbf{E}$  jest macierzą jednostkową trzeciego stopnia.

Nowe pędy  $\mathbf{P} = \text{col}(\mathbf{R}, \mathbf{R}_0)$  znajdziemy zgodnie z Twierdzeniem 5 jako  $\mathbf{P} = \mathbf{A}^{-\text{T}} \mathbf{Q}$ .

Odwrócenie  $\mathbf{A}$  najłatwiej zrealizować pośrednio, rozwiązując prosty układ równań (3.34), co prowadzi do

$$\begin{aligned} \mathbf{r}_1 &= (m - 1) \mathbf{r} + \mathbf{r}_0, \\ \mathbf{r}_2 &= m \mathbf{r} + \mathbf{r}_0, \end{aligned} \quad (3.35)$$

a więc

$$\mathbf{A}^{-\text{T}} = \begin{pmatrix} (m-1)\mathbf{E} & m\mathbf{E} \\ \mathbf{E} & \mathbf{E} \end{pmatrix},$$

i nowe pędy to

$$\begin{aligned} \mathbf{R} &= (m-1)\mathbf{R}_1 + m\mathbf{R}_2, \\ \mathbf{R}_0 &= \mathbf{R}_1 + \mathbf{R}_2. \end{aligned} \quad (3.36)$$

Interpretacja nowych pędów jest prosta:  $\mathbf{R}_0$  to mechaniczny pęd całkowity układu dwóch ciał, będący zarazem pędem ich środka masy. Natomiast

$$\mathbf{R} = \frac{m_1 m_2}{m_1 + m_2} (\dot{\mathbf{r}}_2 - \dot{\mathbf{r}}_1),$$

jest – z dokładnością do mnożnika zależnego od mas – prędkością ciała  $m_2$  względem  $m_1$ .

Wobec braku reszty  $\mathcal{R}$ , nowa funkcja Hamiltona  $\mathcal{K}$  powstaje przez zwykłe podstawienie do  $\mathcal{H}$ . Będziemy do tego potrzebować zależności starych pędów od nowych. W tym celu albo rozwiązujemy układ (3.36), albo korzystamy z  $\mathbf{Q} = \mathbf{A}^T \mathbf{P}$ , co prowadzi do

$$\begin{aligned} \mathbf{R}_1 &= -\mathbf{R} + m\mathbf{R}_0, \\ \mathbf{R}_2 &= \mathbf{R} + (1-m)\mathbf{R}_0. \end{aligned} \quad (3.37)$$

Podstawiając do starej funkcji Hamiltona (2.38)

$$\mathcal{H} = \frac{R_1^2}{2m_1} + \frac{R_2^2}{2m_2} - \frac{k^2 m_1 m_2}{\|\mathbf{r}_2 - \mathbf{r}_1\|},$$

równania (3.34) i (3.37) otrzymamy nowy hamiltonian

$$\begin{aligned} \mathcal{K} &= \frac{1}{2m_1} \left( -\mathbf{R} + \frac{m_1 \mathbf{R}_0}{m_1 + m_2} \right)^2 + \frac{1}{2m_2} \left( \mathbf{R} + \frac{m_2 \mathbf{R}_0}{m_1 + m_2} \right)^2 - \frac{k^2 m_1 m_2}{\|\mathbf{r}\|} = \\ &= \frac{R_0^2}{2(m_1 + m_2)} + \frac{m_1 + m_2}{2m_1 m_2} R^2 - \frac{k^2 m_1 m_2}{r}. \end{aligned} \quad (3.38)$$

Zauważmy, że nastąpiła separacja zagadnienia: hamiltonian  $\mathcal{K}$  jest sumą dwóch wyrazów

$$\mathcal{K}(\mathbf{r}_0, \mathbf{r}, \mathbf{R}_0, \mathbf{R}) = \mathcal{K}_0(\mathbf{r}_0, \mathbf{R}_0) + \mathcal{K}_1(\mathbf{r}, \mathbf{R}),$$

z których każdy jest pełnoprawną funkcją Hamiltona dla swojego podzbioru zmiennych, gdyż równania ruchu środka masy

$$\dot{\mathbf{r}}_0 = \nabla_{\mathbf{R}_0} \mathcal{K} = \nabla_{\mathbf{R}_0} \mathcal{K}_0, \quad \dot{\mathbf{R}}_0 = -\nabla_{\mathbf{r}_0} \mathcal{K} = -\nabla_{\mathbf{r}_0} \mathcal{K}_0,$$

oraz równania ruchu względnego

$$\dot{\mathbf{r}} = \nabla_{\mathbf{R}} \mathcal{K} = \nabla_{\mathbf{R}} \mathcal{K}_1, \quad \dot{\mathbf{R}} = -\nabla_{\mathbf{r}} \mathcal{K} = -\nabla_{\mathbf{r}} \mathcal{K}_1,$$

mają postać kanoniczną.

To oznacza, przy okazji, że oba wyrazy są całkami ruchu na mocy Twierdzenia 3:  $\mathcal{K}_0(\mathbf{R}_0) = \text{const}$ , oraz  $\mathcal{K}_1(\mathbf{r}, \mathbf{R}) = \text{const}$ .

Widzimy, że ruch względny, opisany wektorami współrzędnych  $\mathbf{r}$  i pędu  $\mathbf{R}$ , można rozpatrywać niezależnie od  $\mathbf{r}_0$  i  $\mathbf{R}_0$ . Powstaje układ kanoniczny o 3 stopniach swobody z funkcją Hamiltona

$$\mathcal{K}_1(\mathbf{r}, \mathbf{R}) = \frac{m_1 + m_2}{2 m_1 m_2} R^2 - \frac{k^2 m_1 m_2}{r}, \quad (3.39)$$

i z równaniami ruchu

$$\dot{\mathbf{r}} = \frac{m_1 + m_2}{m_1 m_2} \mathbf{R}, \quad \dot{\mathbf{R}} = -\frac{k^2 m_1 m_2}{r^3} \mathbf{r},$$

który można sprowadzić do znanej postaci

$$\ddot{\mathbf{r}} = -\frac{k^2(m_1 + m_2)}{r^3} \mathbf{r} = -\frac{\mu}{r^3} \mathbf{r}. \quad (3.40)$$

Natomiast hamiltonian  $\mathcal{K}_0(-, \mathbf{R}_0)$ , jako niezależny od położenia środka masy, generuje sześć całek barycentrum. Trzy z nich, to  $\mathbf{R}_0 = \text{const}$  (z Twierdzenia 4), a następne 3 powstają z całkowania równań

$$\dot{\mathbf{r}}_0 = \nabla_{\mathbf{R}_0} \mathcal{K}_0 = \frac{\mathbf{R}_0}{m_1 + m_2} = \text{const}.$$

Idąc dalej, możemy dla zagadnienia względnego (ruch jednej cząstki o masie  $m_2$ ) przeprowadzić skalowanie opisane w rozdziale 2.5.3. Tym razem, zamiast przez masę, podzielimy pęd i hamiltonian przez

$$\alpha = \frac{m_1 m_2}{m_1 + m_2},$$

co doprowadzi do równoważnych, kanonicznych równań ruchu. Nowy pęd

$$\mathbf{R}' = \frac{\mathbf{R}}{\alpha} = \dot{\mathbf{r}}_2 - \dot{\mathbf{r}}_1,$$

jest po prostu prędkością względną. Skalując Hamiltonian, otrzymamy

$$\mathcal{K}'_1 = \frac{\mathcal{K}_1}{\alpha} = \frac{\mathbf{R}' \cdot \mathbf{R}'}{2} - \frac{k^2(m_1 + m_2)}{r}, \quad (3.41)$$

z równaniami kanonicznymi

$$\dot{\mathbf{r}} = \nabla_{\mathbf{R}'} \mathcal{K}'_1 = \mathbf{R}', \quad \dot{\mathbf{R}}' = -\nabla_{\mathbf{r}} \mathcal{K}'_1 = -\frac{k^2(m_1 + m_2)}{r^3} \mathbf{r}, \quad (3.42)$$

które jeszcze łatwiej przekształcić do postaci (3.40).

W ten sposób przeprowadziliśmy redukcję zagadnienia dwóch ciał o 6 stopniach swobody do zagadnienia względnego o 3 stopniach swobody, generując przy okazji 6 całek barycentrum wynikających z osobnego podukładu. W dalszej części wykładu przeprowadzimy kolejne redukcje zagadnienia względnego metodą transformacji kanonicznych, generując kolejne całki ruchu tego zagadnienia. Ostatnim etapem będzie dojście do funkcji Hamiltona zależnej tylko od jednej zmiennej.

### 3.2.5 Zmienne Hilla-Whittakera

Wprowadzimy teraz zmienne kanoniczne, które znacząco upraszczają funkcję Hamiltona zagadnienia dwóch ciał. Są to tak zwane zmienne „biegunowo-węzłowe” lub zmienne Hilla-Whittakera. Przy okazji pokażemy, że ideę transformacji Mathieu (zachowanie wartości formy Pfaffa) można stosować nawet w bardziej złożonej sytuacji niż transformacja punktowa.

Punktem wyjścia będzie względne zagadnienie dwóch ciał (3.41), gdzie  $\mu = k^2(m_1 + m_2)$ , z dodatkowym potencjałem  $V_1$

$$\mathcal{H}(\mathbf{r}, \mathbf{R}', t) = \frac{\mathbf{R}' \cdot \mathbf{R}'}{2} - \frac{\mu}{\|\mathbf{r}\|} + V_1(\mathbf{r}, t), \quad (3.43)$$

opisane przy pomocy kartezjańskiego wektora położenia  $\mathbf{r} = (x, y, z)^T$ , oraz sprzężonego z nim pędu (na jednostkę masy)  $\mathbf{R}' = (X, Y, Z)^T$ . Tym zmien- nym odpowiada kanoniczna jednoforma Pfaffa

$$\Phi = -\mathcal{H} dt + \mathbf{R}' \cdot d\mathbf{r}.$$

Rozbijmy wektor położenia  $\mathbf{r}$  na długość i wersor  $\mathbf{r} = r\hat{\mathbf{r}}$ . Będziemy wtedy mieli

$$d\mathbf{r} = r d\hat{\mathbf{r}} + (dr) \hat{\mathbf{r}}.$$

Różniczkę wersora  $d\hat{\mathbf{r}}$  możemy powiązać z różniczkami kątów Eulera, które zadają orientację bazy radialnej. Jest to zabieg znany już z Rozdziału 1.5.7.

Zastosujemy Twierdzenie 1 dla kątów Eulera 3-1-3:  $\nu$  – długość węzła wstępującego (odstępujemy od zwyczajowego symbolu  $\Omega$ ),  $I$  – nachylenie orbity, oraz  $\vartheta$  – argument szerokości (suma argumentu perycentrum i anomalii prawdziwej). Zamiast prędkości będziemy używać różniczek (mnożymy obie strony przez  $dt$ ), co prowadzi do

$$d\mathbf{r} = r \left( \hat{\mathbf{z}}d\nu + \hat{\mathbf{m}}dI + \hat{\mathbf{G}}d\vartheta \right) \times \hat{\mathbf{r}} + (dr) \hat{\mathbf{r}}. \quad (3.44)$$

Wersory  $\hat{\mathbf{z}}$ ,  $\hat{\mathbf{m}}$ , i  $\hat{\mathbf{G}}$  skierowane są odpowiednio wzdłuż osi  $Oz$  przyjętego układu, do węzła wstępującego orbity (1.49) i wzdłuż wektora momentu pędu

$$\mathbf{G} = \mathbf{r} \times \mathbf{R}'. \quad (3.45)$$

W przytoczonej wyżej formie Pfaffa możemy więc wykonać podstawienie

$$\begin{aligned} \mathbf{R}' \cdot d\mathbf{r} &= \mathbf{R}' \cdot \left( \hat{\mathbf{z}}d\nu + \hat{\mathbf{m}}dI + \hat{\mathbf{G}}d\vartheta \right) \times \mathbf{r} + \mathbf{R}' \cdot \hat{\mathbf{r}} dr = \\ &= \mathbf{R}' \cdot \hat{\mathbf{r}} dr + \mathbf{R}' \cdot (\hat{\mathbf{z}} \times \mathbf{r}) d\nu \\ &\quad + \mathbf{R}' \cdot (\hat{\mathbf{m}} \times \mathbf{r}) dI + \mathbf{R}' \cdot (\hat{\mathbf{G}} \times \mathbf{r}) d\vartheta = \\ &= \mathbf{R}' \cdot \hat{\mathbf{r}} dr + \hat{\mathbf{z}} \cdot \mathbf{G} d\nu + \hat{\mathbf{G}} \cdot \mathbf{G} d\vartheta, \end{aligned} \quad (3.46)$$

gdzie po skorzystaniu z cyklicznej zamiany czynników w iloczynie mieszanym podstawiliśmy (3.45). Różniczka  $dI$  znikła, gdyż moment pędu jest prostopadły do linii węzłów; w tej sytuacji zostały nam w formie Pfaffa trzy różniczki, związane ze zmiennymi  $r$ ,  $\nu$  i  $\vartheta$ . Transformacja  $(x, y, z) \rightarrow (r, \vartheta, \nu)$ , nie jest transformacją punktową, ale dopóki w formie Pfaffa nie pojawiły się różniczki pędów, warunek (3.17) pozostaje w mocy.

Ponieważ transformacja nie jest jawnie zależna od czasu, możemy przyjąć, że trzy nowe współrzędne uogólnione

$$\mathbf{p} = (r, \vartheta, \nu)^T,$$

i sprzężone z nimi pędy

$$\mathbf{P} = (R, \Theta, N)^T,$$

oraz hamiltonian  $\mathcal{K}$  są zdefiniowane warunkiem (3.17)

$$-\mathcal{H}dt + Xdx + Ydy + Zdz = -\mathcal{K}dt + Rdr + \Theta d\vartheta + N d\nu,$$

co, po uwzględnieniu (3.46), prowadzi do  $\mathcal{K} = \mathcal{H}$ , oraz

$$\begin{cases} r, & R = \mathbf{R}' \cdot \hat{\mathbf{r}} = \dot{r}, \\ \vartheta = f + \omega, & \Theta = \hat{\mathbf{G}} \cdot \mathbf{G} = G = \sqrt{\mu p}, \\ \nu = \Omega, & N = \hat{\mathbf{z}} \cdot \mathbf{G} = G \cos I. \end{cases} \quad (3.47)$$

Użyliśmy do zdefiniowania współrzędnych  $(r, \vartheta, \nu)$  i sprzężonych z nimi pędów  $(R, \Theta, N)$  symboli znanych z zagadnienia dwóch ciał: anomalii prawdziwej  $f$ , argumentu perycentrum  $\omega$ , parametru (*semilatus*) orbity  $p$ , długości węzła wstępującego  $\Omega$  i nachylenia  $I$ . Zauważmy, że pęd  $\Theta$  to nic innego jak długość wektora momentu pędu (stała pól), zaś  $N$  to rzut wektora momentu pędu na oś biegunową  $Oz$  kartezjańskiego układu współrzędnych.

Wyrażając funkcję Hamiltona przy użyciu nowych zmiennych, możemy zauważyć, że prędkość całkowita jest zawsze sumą prędkości radialnej i transwersalnej, więc

$$\mathbf{R}' \cdot \mathbf{R}' = X^2 + Y^2 + Z^2 = R^2 + (\Theta/r)^2.$$

Jeśli we współrzędnych kartezjańskich potencjał jest sumą keplerowskiego

$$V_0 = -\frac{\mu}{\sqrt{x^2 + y^2 + z^2}},$$

i dowolnego  $V_1(x, y, z, t)$ , to funkcja Hamiltona w zmiennych Hilla-Whittakera ma postać

$$\mathcal{K} = \frac{R^2}{2} + \frac{\Theta^2}{2r^2} - \frac{\mu}{r} + V_1(r, \vartheta, \nu, -, \Theta, N, t), \quad (3.48)$$

gdzie zależność  $V_1$  od pędów wynika stąd, że współrzędne kartezjańskie można wyrazić przez  $r, \vartheta, \nu$  tylko w powiązaniu z nachyleniem  $I = \arccos N/\Theta$ . Jak widać, wykroczyliśmy tutaj poza pierwotne ograniczenie (3.16).

Na zakończenie wypiszmy równania kanoniczne Hamiltona

$$\begin{aligned} \dot{r} &= \frac{\partial \mathcal{K}}{\partial R} = R, & \dot{R} &= -\frac{\partial \mathcal{K}}{\partial r} = \frac{\Theta^2}{r^3} - \frac{\mu}{r^2} - \frac{\partial V_1}{\partial r}, \\ \dot{\vartheta} &= \frac{\partial \mathcal{K}}{\partial \Theta} = \frac{\Theta}{r^2} + \frac{\partial V_1}{\partial \Theta}, & \dot{\Theta} &= -\frac{\partial \mathcal{K}}{\partial \vartheta} = -\frac{\partial V_1}{\partial \vartheta}, \\ \dot{\nu} &= \frac{\partial \mathcal{K}}{\partial N} = \frac{\partial V_1}{\partial N}, & \dot{N} &= -\frac{\partial \mathcal{K}}{\partial \nu} = -\frac{\partial V_1}{\partial \nu}, \end{aligned} \quad (3.49)$$

Zuważmy, że gdy  $V_1 = 0$ , czyli w zagadnieniu dwóch ciał, układ (3.49) rozpada się na dwie części: równania dla  $R, r$  rozwiązujemy niezależnie od pozostałych, gdyż  $\Theta$ , podobnie jak  $N$  i  $\nu$  jest całką ruchu. Z pozostałych czterech równań tylko równanie dla  $\vartheta$  jest nietrywialne, ale po znalezieniu zależności  $r(t)$  daje się całkować przez kwadratury. Warto się też zastanowić jakie kryteria musi spełniać zaburzenie  $V_1$ , aby zagadnienie pozostawało całkowne...

Zmienne Hilla-Whittakera stanowią z jednej strony ważny krok do rozwiązania zagadnienia dwóch ciał w ramach formalizmu kanonicznego, a z

drugiej strony są bardzo wygodne w wielu zagadnieniach ruchu zaburzonego. Jedyłą ich słabością jest możliwość wystąpienia osobliwości dla małych nachyleń, ale ten problem można obejść drogą dość prostej modyfikacji polegającej na wprowadzeniu kąta  $\vartheta + \nu$ .

### 3.3 Transformacje zadane funkcją tworzącą

#### 3.3.1 Funkcja tworząca zmiennych mieszanych

Wspomnieliśmy już, że transformacja punktowa zakłada dość mocne ograniczenia i tylko pewne przypadki transformacji kanonicznych można realizować przy jej pomocy. Z drugiej strony, kryteria kanoniczności z wcześniejszych rozdziałów nie są konstruktywne i pozwalają jedynie sprawdzić, czy zadana transformacja jest kanoniczna czy nie. Obecnie zajmiemy się podejściem zarówno konstruktywnym jak i dostatecznie ogólnym: omówimy transformacje kanoniczne zadane funkcją tworzącą zmiennych mieszanych. Nazwa „funkcja tworząca zmiennych mieszanych” wynika z faktu, że zawsze mamy w niej połowę starych i połowę nowych zmiennych.

**TWIERDZENIE 6** Niech  $\xi = \text{col}(\mathbf{q}, \mathbf{Q})$  i  $\eta = \text{col}(\mathbf{p}, \mathbf{P})$ . Transformacja  $\xi \leftrightarrow \eta$  jest kanoniczna, jeśli istnieje przynajmniej jeden z czterech rodzajów *funkcji tworzącej zmiennych mieszanych*  $F_k$ , spełniający odpowiednio warunki:

$$F_1(\mathbf{q}, \mathbf{p}, t) : \quad Q_i = \frac{\partial F_1}{\partial q_i}, \quad P_i = -\frac{\partial F_1}{\partial p_i}, \quad (3.50)$$

$$F_2(\mathbf{q}, \mathbf{P}, t) : \quad Q_i = \frac{\partial F_2}{\partial q_i}, \quad p_i = \frac{\partial F_2}{\partial P_i}, \quad (3.51)$$

$$F_3(\mathbf{Q}, \mathbf{p}, t) : \quad q_i = -\frac{\partial F_3}{\partial Q_i}, \quad P_i = -\frac{\partial F_3}{\partial p_i}, \quad (3.52)$$

$$F_4(\mathbf{Q}, \mathbf{P}, t) : \quad q_i = -\frac{\partial F_4}{\partial Q_i}, \quad p_i = \frac{\partial F_4}{\partial P_i}, \quad (3.53)$$

gdzie indeks  $i$  przebiega stopnie swobody od 1 do  $M$ . Dla każdego typu funkcji tworzącej  $F_k$ , nowy Hamiltonian  $\mathcal{K}$  jest dany równaniem

$$\mathcal{K}(\mathbf{p}, \mathbf{P}, t) = \mathcal{H}(\xi(\eta, t), t) + \mathcal{R}, \quad \text{gdzie} \quad \mathcal{R} = \frac{\partial F_k}{\partial t}. \quad (3.54)$$

Dowód tego twierdzenia jest prosty, jeśli odwołamy się do Twierdzenia 2. Weźmy kanoniczną jednoformę (2.24)

$$\Phi = -\mathcal{H} dt + \mathbf{Q} \cdot d\mathbf{q}.$$



Jeśli dodamy do niej różniczkę zupełną dowolnej funkcji  $dS$  i pomnożymy przez dowolną stałą, to będzie ona generować równania ruchu tego samego układu. Niech  $S = -F_2$ ; dodajemy do  $\Phi$  różniczkę  $d(-F_2(\mathbf{q}, \mathbf{P}, t))$ , co prowadzi do nowej jednoformy

$$\begin{aligned}\Phi' &= c[\Phi + d(-F_2)] = \\ &= c[-\mathcal{H} dt + \mathbf{Q} \cdot d\mathbf{q} - \nabla_{\mathbf{q}} F_2 \cdot d\mathbf{q} - \nabla_{\mathbf{P}} F_2 \cdot d\mathbf{P} - D_t F_2 dt] = \\ &= c[-(\mathcal{H} + D_t F_2) dt + (\mathbf{Q} - \nabla_{\mathbf{q}} F_2) \cdot d\mathbf{q} - \nabla_{\mathbf{P}} F_2 \cdot d\mathbf{P}].\end{aligned}$$

Widzimy, że jeśli zgodnie z równaniami (3.51) i (3.54) przyjmiemy

$$\nabla_{\mathbf{q}} F_2 = \mathbf{Q}, \quad \nabla_{\mathbf{P}} F_2 = \mathbf{p}, \quad \mathcal{R} = D_t F_2, \quad \text{i} \quad c = -1,$$

to powstanie jednoforma kanoniczna typu (2.25)

$$\Phi' = \mathcal{K} dt + \mathbf{p} \cdot d\mathbf{P},$$

która generuje równania Hamiltona w zmiennych  $\mathbf{p}, \mathbf{P}$  z funkcją  $\mathcal{K} = \mathcal{H} + \mathcal{R}$ .

Dowód dla pozostałych typów  $F_j$  przebiega podobnie:

- Dla  $F_1$  dodajemy do (2.24) różniczkę funkcji  $S = -F_1$  i z mnożnikiem  $c = 1$  otrzymamy  $\Phi' = -\mathcal{K} dt + \mathbf{P} \cdot d\mathbf{p}$ .
- Dla  $F_3$  dodajemy do (2.25) różniczkę funkcji  $S = F_3$  i dla  $c = -1$  dochodzimy do tego samego  $\Phi'$ , jak w przypadku  $F_1$ .
- Dla  $F_4$  dodajemy do (2.25) różniczkę  $dF_4$  i przyjmując  $c = 1$  mamy  $\Phi' = \mathcal{K} dt + \mathbf{p} \cdot d\mathbf{P}$ , jak dla  $F_2$ .

### 3.3.2 Transformacja tożsamościowa

Rozpatrzmy przykład trywialny, ale ważny i użyteczny – transformację tożsamościową, która nie zmienia zmiennych, to znaczy

$$\boldsymbol{\eta} = \boldsymbol{\xi}, \quad \text{czyli} \quad \mathbf{p} = \mathbf{q}, \quad \mathbf{P} = \mathbf{Q}.$$

Transformacja tożsamościowa ma funkcje tworzące typu  $F_2$  i  $F_3$ . Są to

$$F_2(\mathbf{q}, \mathbf{P}) = \mathbf{q}^T \mathbf{P} = q_1 P_1 + \dots + q_M P_M, \quad (3.55)$$

gdyż istotnie

$$Q_i = \frac{\partial F_2}{\partial q_i} = P_i, \quad p_i = \frac{\partial F_2}{\partial P_i} = q_i,$$

oraz

$$F_3(\mathbf{Q}, \mathbf{p}) = -\mathbf{Q}^T \mathbf{p} = -Q_1 p_1 - \dots - Q_M p_M, \quad (3.56)$$

gdź wtedy

$$q_i = -\frac{\partial F_3}{\partial Q_i} = p_i, \quad P_i = -\frac{\partial F_3}{\partial p_i} = Q_i.$$

Zauważmy, że nie można podać funkcji typu  $F_1$  lub  $F_4$  dla transformacji tożsamościowej.

### 3.3.3 Zmienne ką-t-działanie

Rozpatrzmy zagadnienie o  $M$  stopniach swobody, z funkcją Hamiltona  $\mathcal{H}(\mathbf{q}, \mathbf{Q})$ , które potrafimy rozwiązać w sposób ścisły i rozwiązanie ma charakter ograniczonych oscylacji. W zagadnieniu takim można wprowadzić szczególny rodzaj zmiennych kanonicznych nazywanych **zmiennymi ką-t-działanie**  $\psi$ ,  $\Psi$ . Mają one następujące własności:

1. Nowa funkcja Hamiltona  $\mathcal{K}(-, \Psi)$  zależy tylko od pędów  $\Psi$ , czyli zmiennych typu działanie.
2. Zmienne cykliczne typu kąt  $\psi$  mają tę własność, że kiedy do każdej z nich dodamy  $2\pi$ , to otrzymamy tę samą konfigurację w przestrzeni fazowej. Inaczej mówiąc, jeśli  $\xi = \text{col}(\mathbf{q}, \mathbf{Q})$ , to

$$\xi(\psi_1, \dots, \psi_M, \Psi_1, \dots, \Psi_M) = \xi(\psi_1 + 2\pi, \dots, \psi_M + 2\pi, \Psi_1, \dots, \Psi_M). \quad (3.57)$$

Równania kanoniczne dla zmiennych ką-t-działanie mają prostą postać

$$\dot{\psi}_i = \frac{\partial \mathcal{K}(-, \Psi)}{\partial \Psi_i} = \omega_i(\Psi) = \text{const}, \quad \dot{\Psi}_i = -\frac{\partial \mathcal{K}(-, \Psi)}{\partial \psi_i} = 0, \quad (3.58)$$

z trywialnym rozwiązaniem

$$\psi_i = \omega_i(t - t_0) + \psi_i(t_0), \quad \Psi_i = \text{const}. \quad (3.59)$$

A zatem zmienne typu działanie są stałe, natomiast zmienne typu kąt są liniowymi funkcjami czasu ze stałymi częstotliwościami zależnymi od działań  $\Psi$ . Jak widać, znalezienie zmiennych typu kąt działanie dla danego zagadnienia jest równoznaczne z jego rozwiązaniem. Transformacji kanonicznej, która prowadzi do zmiennych ką-t-działanie poszukujemy zazwyczaj przy użyciu funkcji tworzącej.

### Oscylator harmoniczny: zmienne ką-t-działanie

W rozdziale 3.1.3 podaliśmy transformację Poincarégo, w wyniku której hamiltonian oscylatora harmonicznego (3.11)

$$\mathcal{H} = \frac{1}{2} (X^2 + \omega^2 x^2),$$

przyjmuje postać (3.13)

$$\mathcal{K} = \omega L.$$

Innymi słowy, transformacja  $\xi = (x, X)^T \leftrightarrow \eta = (\ell, L)^T$ , zdefiniowana poprzez (3.12), wprowadza zmienne ką-t-działanie dla tego zagadnienia. Spróbujmy teraz wyprowadzić wzory (3.12) zamiast przyjmować je jako dane z góry. Będzie to więc rozwiązanie problemu całkowania równań ruchu oscylatora harmonicznego drogą redukcji hamiltonianu z pierwotnej postaci  $\mathcal{H}$  do założonej  $\mathcal{K}$ .

Poszukiwana transformacja musi być kanoniczna, gdyż tylko wtedy ma sens operowanie pojęciem nowego hamiltonianu. Transformacja będzie kanoniczna, jeśli od samego początku przyjmujemy, że posiada funkcję tworzącą – na przykład typu  $F_2$ . Mamy wtedy

$$X = \frac{\partial F_2(x, L)}{\partial x}, \quad \ell = \frac{\partial F_2(x, L)}{\partial L}, \quad (3.60)$$

choć jeszcze nie wiemy jaka jest postać  $F_2$ . Co ciekawe, wiedza na temat jawnej postaci  $F_2$  nie będzie nam wcale potrzebna...

Ponieważ  $F_2$  nie zależy jawnie od czasu, to  $\mathcal{H} = \mathcal{K}$  i podstawiając (3.60) możemy napisać

$$\frac{1}{2} \left( \frac{\partial F_2}{\partial x} \right)^2 + \frac{\omega^2 x^2}{2} = \omega L,$$

a więc, zaniedbując kwestię znaku,

$$X = \frac{\partial F_2}{\partial x} = \sqrt{2\omega L - \omega^2 x^2}, \quad (3.61)$$

Jak widać, równanie (3.61) dopuszcza znalezienie  $F_2$  drogą całkowania, a ponieważ po jego lewej stronie stał znak pochodnej cząstkowej  $\partial x$ , to całkowanie prowadzimy traktując  $x$  jako jedyną zmienną

$$F_2 = \int \sqrt{2\omega L - \omega^2 x^2} dx. \quad (3.62)$$

Pozostaje problem stałej addytywnej całkowania, czy też – co na jedno wychodzi – granic całkowania. Jako wygodną konwencję przyjmijmy, że dolną

granicą całkowania jest  $x = 0$ . Mówiąc mniej ściśle, całkujemy od maksimum  $X$ . Górną granicą jest „bieżąca wartość”  $x$ .

Równanie (3.62) definiuje nam pośrednio funkcję  $F_2$ . Aby znaleźć nową zmienną  $\ell$ , sięgamy do drugiego równań (3.60)

$$\ell = \frac{\partial F_2}{\partial L} = \frac{\partial}{\partial L} \int \sqrt{2\omega L - \omega^2 x^2} dx = \int \frac{\omega dx}{\sqrt{2\omega L - \omega^2 x^2}}.$$

Różniczkowanie pod znakiem całki było możliwe tylko dla tego, że  $L$  nie jest traktowane jako zmienna (por. 3.61 !), lecz jako parametr. W ten sposób dochodzimy do definicji kąta  $\ell$

$$\ell = \int_0^x \frac{dx}{\sqrt{\frac{2L}{\omega} - x^2}} = \arcsin \left( \frac{x}{\sqrt{2L/\omega}} \right).$$

W świetle poprzedniego równania i wzoru (3.61) dochodzimy do zgodnego z równaniami (3.12) wniosku

$$\begin{aligned} x &= \sqrt{2L/\omega} \sin \ell, \\ X &= \sqrt{2\omega L - \omega^2 x^2} = \sqrt{2\omega L} \cos \ell. \end{aligned} \tag{3.63}$$

Kanoniczności tej transformacji nie musimy już sprawdzać, gdyż wynika ona z faktu istnienia funkcji  $F_2(x, L)$ ; dokładna postać tej funkcji okazała się nam niepotrzebna.

### 3.3.4 Zmienne Delaunaya

Wprowadzimy teraz zmienne ką-działanie dla względnego zagadnienia dwóch ciał. Jak dotąd, doprowadziliśmy hamiltonian tego zagadnienia do stosunkowo prostej postaci (3.48)

$$\mathcal{H}(r, R, \Theta) = \frac{1}{2} \left( R^2 + \frac{\Theta^2}{r^2} \right) - \frac{\mu}{r}.$$

Było to możliwe dzięki użyciu zmiennych Hilla-Whittakera (3.47). Naszym zadaniem będzie teraz znalezienie zmiennych typu ką-działanie, dla których funkcja Hamiltona zagadnienia dwóch ciał zależęć będzie tylko od jednego pędu. Chcemy więc znaleźć transformację kanoniczną

$$\xi = (r, \vartheta, \nu, R, \Theta, N)^T \leftrightarrow \eta = (\ell, g, h, L, G, H)^T,$$

gdzie zmienne  $\eta$  noszą nazwę zmiennych Delaunaya. Aby transformacja była kanoniczna, założymy dla niej funkcję tworzącą typu  $F_2$

$$F_2 = F_2(r, \vartheta, \nu, L, G, H).$$

Ponieważ  $\frac{\partial F_2}{\partial t} = 0$ , to transformacja nie będzie posiadała reszty i nowy hamiltonian  $\mathcal{K}$  będzie dany jako

$$\mathcal{K}(L) = \mathcal{H}(r, R, \Theta).$$

Zacznijmy od pary zmiennych, które nie wymagają żadnej modyfikacji. Są to długość węzła wstępującego  $\nu$  i sprzężona z nim kanonicznie składowa z momentu pędu  $N$ . Żadna z nich nie występuje w  $\mathcal{H}$ , a więc nie pojawi się i w  $\mathcal{K}$ ; wystarczy przypisać do nich pierwszą parę zmiennych Delaunaya

$$h = \nu, \quad H = N.$$

Taka transformacja tożsamościowa jest już nam znana z Rozdziału 3.3.2. Możemy więc rozbić funkcję tworzącą  $F_2$  na dwie części

$$F_2(r, \vartheta, \nu, L, G, H) = \nu H + F_2^*(r, \vartheta, L, G). \quad (3.64)$$

Pierwszy wyraz prawej strony definiuje transformację tożsamościową  $h = \nu$ ,  $H = N$ , oczywiście pod warunkiem, że  $F_2^*$  nie zależy od  $\nu$  i  $H$ .

Kolejną zmienną, która może pozostać bez zmian, jest pęd uogólniony  $\Theta$ , czyli całkowity moment pędu. Tym razem jednak, o ile można przyjąć  $G = \Theta$ , to nie można utożsamić  $g$  i  $\vartheta$ . Wiemy bowiem, że

$$\dot{\vartheta} = \frac{\partial \mathcal{H}}{\partial \Theta} = \frac{\Theta}{r^2},$$

natomiast poszukiwany przez nas nowy hamiltonian  $\mathcal{K}$  ma zależeć tylko od  $L$ , co będzie oznaczać

$$\dot{g} = \frac{\partial \mathcal{K}}{\partial G} = 0.$$

Aby więc  $G = \Theta$ , lecz  $g \neq \vartheta$ , przedstawmy  $F_2$  w postaci

$$F_2(r, \vartheta, \nu, L, G, H) = \nu H + \vartheta G + \Psi(r, L, G), \quad (3.65)$$

gdzie nieznana jeszcze funkcja  $\Psi$  modyfikować będzie związek między argumentem szerokości  $\vartheta$  i kątem  $g$ .

Na tym kończymy część elementarną, dotyczącą możliwych transformacji tożsamościowych. Przechodzimy teraz do analizy układu definiowanego funkcją tworzącą (3.65). W myśl równania (3.51) mamy

$$\left\{ \begin{array}{l} \ell = \frac{\partial F_2}{\partial L} = \frac{\partial \Psi}{\partial L}, \\ R = \frac{\partial F_2}{\partial r} = \frac{\partial \Psi}{\partial r}, \\ g = \frac{\partial F_2}{\partial G} = \vartheta + \frac{\partial \Psi}{\partial G}, \\ \Theta = \frac{\partial F_2}{\partial \vartheta} = G, \\ h = \frac{\partial F_2}{\partial H} = \nu, \\ N = \frac{\partial F_2}{\partial \nu} = H. \end{array} \right. \quad (3.66)$$

Ostatnie trzy z powyższych równań zawierają gotowe wyniki, natomiast pierwsze trzy posłużą do zdefiniowania brakujących zmiennych. Zauważmy na początek, że funkcja Hamiltona w tym zagadnieniu spełnia całkę energii (siły żywej)

$$\mathcal{H} = \mathcal{K}(L) = E,$$

przy czym stała energii  $E = E(L)$  może zależeć jedynie od pędu  $L$ . Przyrównując  $\mathcal{H}$  i  $E$ , oraz podstawiając  $\Theta = G$ , otrzymamy wyjściowe równanie

$$\frac{1}{2} \left( R^2 + \frac{G^2}{r^2} \right) - \frac{\mu}{r} = E.$$

Rozwiązując to równanie względem  $R$ , otrzymamy

$$R = \sqrt{2E + \frac{2\mu}{r} - \frac{G^2}{r^2}} = Q(r, G, L), \quad (3.67)$$

gdzie wprowadziliśmy pomocniczy symbol  $Q$ . W myśl pierwszego z równań (3.66), oraz równania (3.67)

$$\frac{\partial \Psi}{\partial r} = Q(r, G, L).$$

A zatem, niewiadoma funkcja  $\Psi$  może zostać przedstawiona jako kwadratura względem  $r$ , przy czym wszystkie pozostałe zmienne traktujemy jak parametry

$$\Psi = \int Q dr = \int \sqrt{2E + \frac{2\mu}{r} - \frac{G^2}{r^2}} dr. \quad (3.68)$$

Pozostaje problem dolnej granicy całkowania. Nauczeni doświadczeniem z anomaliami opisującymi ruch keplerowski, przyjmijmy, że całkowanie odbywa się od perycentrum, czyli punktu gdzie  $R = 0$  a  $r$  przyjmuje wartość minimalną.

Równanie (3.68) stanowi klucz do definicji brakujących zmiennych Delaunaya. Zacznijmy od  $g$ . W myśl trzeciego z wzorów (3.66)

$$g = \vartheta + \frac{\partial \Psi}{\partial G} = \vartheta + \frac{\partial}{\partial G} \int Q dr = \vartheta + \int \frac{\partial Q}{\partial G} dr.$$

Ponieważ

$$\frac{\partial Q}{\partial G} = \frac{1}{2Q} \left( -\frac{2G}{r^2} \right) = -\frac{G}{Qr^2},$$

to

$$g = \vartheta - G \int \frac{dr}{Qr^2}. \quad (3.69)$$

Wykorzystajmy znany z równań w zmiennych Hilla-Whittakera fakt, że  $R = \dot{r}$ , a będziemy mogli napisać

$$\frac{G dr}{Qr^2} = \frac{G dr}{Rr^2} = \frac{G dr}{\frac{dr}{dt} r^2} = \frac{G}{r^2} dt.$$

Wystarczy teraz przypomnieć sobie całkę pól (II prawo Keplera) lub zdefiniować anomalię prawdziwą  $f$  poprzez związek  $\frac{df}{dt} = G/r^2$  a okaże się, że

$$g = \vartheta - \int df = \vartheta - f = \omega. \quad (3.70)$$

Fakt, że całkujemy od perycentrum był tu istotny, gdyż inaczej mielibyśmy  $g = \omega + \text{const}$ .

Znamy już cztery zmienne Delaunaya; aby znaleźć kolejną, przejdźmy do pierwszego z równań (3.66)

$$\ell = \frac{\partial \Psi}{\partial L} = \frac{\partial}{\partial L} \int Q \, dr = \int \frac{\partial Q}{\partial L} \, dr = \int \frac{1}{Q} \frac{\partial E}{\partial L} \, dr.$$

I tym razem wykorzystamy związki  $Q = R$  i  $R = \frac{dr}{dt}$  przy czym, jak zawsze, wolno nam to zrobić dopiero po wykonaniu różniczkowania. W efekcie

$$\ell = \int \frac{1}{R} \frac{\partial E}{\partial L} \, dr = \int \frac{\partial E}{\partial L} \, dt = \frac{\partial E}{\partial L} (t - t_p),$$

gdzie  $t_p$  oznacza epokę przejścia perycentrum. Jeśli teraz zechcemy, aby

$$\frac{\partial E}{\partial L} = n = \sqrt{\frac{\mu}{a^3}},$$

to możemy prosto zinterpretować kąt  $\ell$  jako anomalię średnią

$$\ell = M.$$

Pozostaje już tylko kwestia pędu  $L$ . Jego definicja jest konsekwencją dopiero co dokonanego wyboru, gdyż jeśli

$$\frac{\partial E}{\partial L} = \frac{dE}{dL} = n,$$

to

$$dL = \frac{dE}{n} = \sqrt{\frac{a^3}{\mu}} d\left(-\frac{\mu}{2a}\right) = \sqrt{\frac{a^3}{\mu}} \frac{\mu}{2a^2} da = \frac{\sqrt{\mu}}{2} \frac{da}{\sqrt{a}},$$

i ostatecznie

$$L = \sqrt{\mu a}, \quad a = \frac{L^2}{\mu}.$$

Możemy już podsumować wyniki i podać definicje zmiennych Delaunaya oraz ich związek z elementami keplerowskimi orbity.

$$\begin{aligned} \ell &= M, & L &= \sqrt{\mu a}, \\ g &= \omega, & G &= \sqrt{\mu p} = L \sqrt{1 - e^2}, \\ h &= \Omega, & H &= G_z = G \cos I. \end{aligned} \quad (3.71)$$



Jeśli zaś chodzi o funkcję Hamiltona, to w świetle całki energii  $\mathcal{K} = -\mu/(2a)$  a zatem

$$\mathcal{K} = -\frac{\mu^2}{2L^2}. \quad (3.72)$$

Jest to najprostsza i ostateczna postać funkcji Hamiltona dla względnego zagadnienia dwóch ciał. Zmienne Delaunaya są zmiennymi kąta-działanie dla tego zagadnienia, a zatem ich znalezienie jest równoznaczne z rozwiązaniem problemu ruchu dwóch ciał. Wypisując równania kanoniczne

$$\begin{aligned} \dot{\ell} &= \frac{\partial \mathcal{K}}{\partial L} = \frac{\mu^2}{L^3} = n, & \dot{L} &= -\frac{\partial \mathcal{K}}{\partial \ell} = 0, \\ \dot{g} &= \frac{\partial \mathcal{K}}{\partial G} = 0, & \dot{G} &= -\frac{\partial \mathcal{K}}{\partial g} = 0, \\ \dot{h} &= \frac{\partial \mathcal{K}}{\partial H} = 0, & \dot{H} &= -\frac{\partial \mathcal{K}}{\partial h} = 0, \end{aligned} \quad (3.73)$$

stwierdzamy, że istotnie wszystkie zmienne typu „działanie” ( $L, G, H$ ) są stałe, zaś kąt  $\ell$  jest liniową funkcją czasu. Fakt, iż dodatkowo  $\dot{g} = \dot{h} = 0$  dla wszelkich wartości pędów, oznacza, że w zagadnieniu dwóch ciał występuje tzw. degeneracja właściwa (inaczej: degeneracja istotna). Można powiedzieć, że zagadnienie dwóch ciał posiada efektywnie jeden stopień swobody.

### 3.4 Równania planetarne Lagrange'a

Przez równania planetarne Lagrange'a rozumiemy wariant równań Gaussa dla uzmiennionych elementów keplerowskich, który można stosować dla sił zaburzających posiadających potencjał. Ponieważ wspomnieliśmy już, że dla każdej praktycznie siły potencjalnej można sformułować równania ruchu w postaci kanonicznej, posłużymy się do ich wyprowadzenia formalizmem hamiltonowskim, choć oczywiście same równania planetarne Lagrange'a nie będą kanoniczne. Nazwa „równania planetarne” ma charakter tradycyjny i nie oznacza, bynajmniej, że można ich używać tylko do opisu ruchu planet; została ona wprowadzona dla odróżnienia tych równań innych równań Lagrange'a – na przykład równań Lagrange'a drugiego rodzaju. Do końca tego rozdziału, mówiąc o równaniach Lagrange'a będziemy mieli na myśli „równania planetarne”.

Formalizm kanoniczny pozwala na wyjątkowo zwarte i oszczędne wyprowadzenie równań Lagrange'a pod warunkiem, że znamy takie pojęcia jak zmienne Delaunaya i nawiasy Poissona. Zastosowanie nawiasów Poissona jest o wiele prostsze niż używanie spotykanych w starszych podręcznikach nawiasów Lagrange'a. Aby znaleźć równania ruchu dla uzmiennionych elementów keplerowskich, dalej zwanych zmiennymi keplerowskimi, przyjmijmy funkcję Hamiltona (3.72) z dodatkowym potencjałem zaburzającym  $V(\mathbf{r}, t)$ . Ponieważ położenie  $\mathbf{r}$  jest funkcją wszystkich sześciu zmiennych keplerowskich, mamy

$$\mathcal{H} = -\frac{\mu}{2a} + V(a, e, I, M, \omega, \Omega, t). \quad (3.74)$$

Dla każdej ze zmiennych keplerowskich  $\alpha \in \{a, e, I, M, \omega, \Omega\}$ , możemy posłużyć się wzorem (2.29)

$$\dot{\alpha} = \{ \alpha, \mathcal{H} \}, \quad (3.75)$$

gdzie nawias Poissona (2.27) będziemy liczyć względem najbardziej wygodnego w tym kontekście zestawu zmiennych kanonicznych – zmiennych Delaunaya (3.71), to znaczy

$$\{ \alpha, \mathcal{H} \} = \frac{\partial \alpha}{\partial \ell} \frac{\partial \mathcal{H}}{\partial L} - \frac{\partial \alpha}{\partial L} \frac{\partial \mathcal{H}}{\partial \ell} + \frac{\partial \alpha}{\partial g} \frac{\partial \mathcal{H}}{\partial G} - \frac{\partial \alpha}{\partial G} \frac{\partial \mathcal{H}}{\partial g} + \frac{\partial \alpha}{\partial h} \frac{\partial \mathcal{H}}{\partial H} - \frac{\partial \alpha}{\partial H} \frac{\partial \mathcal{H}}{\partial h}. \quad (3.76)$$

Jak widać, musimy tylko znaleźć wzory dla pochodnych cząstkowych zmiennych keplerowskich względem zmiennych Delaunaya a reszta procedury będzie już elementarna.

### 3.4.1 Pochodne cząstkowe zmiennych keplerowskich względem zmiennych Delaunaya

Ponieważ zmienne Delaunaya zdefiniowaliśmy jako proste funkcje zmiennych keplerowskich (3.71), wystarczy zróżniczkować wprowadzone związki.

**Kąty  $M$ ,  $\omega$  i  $\Omega$**

Jeśli chodzi o wymienione wyżej kąty, to sprawa jest prosta, gdyż każdy z nich jest w istocie zmienną Delaunaya. Mamy więc dla dowolnej funkcji  $\phi(M, \omega, \Omega)$

$$\frac{\partial \phi}{\partial \ell} = \frac{\partial \phi}{\partial M}, \quad \frac{\partial \phi}{\partial g} = \frac{\partial \phi}{\partial \omega}, \quad \frac{\partial \phi}{\partial h} = \frac{\partial \phi}{\partial \Omega}, \quad (3.77)$$

**Półoś wielka, mimośród, nachylenie**

Rozpatrzmy układ równań definiujących pędy Delaunaya (3.71)

$$\begin{cases} L = \sqrt{\mu a}, \\ G = L \sqrt{1 - e^2}, \\ H = G \cos I \end{cases}$$

Różniczkując obie strony tego układu otrzymamy

$$\begin{cases} dL = \sqrt{\frac{\mu}{a}} \frac{da}{2}, \\ dG = dL \sqrt{1 - e^2} - \frac{Le}{\sqrt{1 - e^2}} de, \\ dH = dG \cos I - G \sin I dI. \end{cases}$$

Wystarczy teraz rozwiązać ten układ względem  $da$ ,  $de$  oraz  $dI$ , aby otrzymać poszukiwane pochodne cząstkowe.

$$\begin{cases} da = 2 \sqrt{\frac{a}{\mu}} dL = \frac{\partial a}{\partial L} dL, \\ de = \frac{1 - e^2}{Le} dL - \frac{\sqrt{1 - e^2}}{Le} dG = \frac{\partial e}{\partial L} dL + \frac{\partial e}{\partial G} dG, \\ dI = \frac{e dG}{G \sin I} - \frac{dH}{G \sin I} = \frac{\partial I}{\partial G} dG + \frac{\partial I}{\partial H} dH. \end{cases} \quad (3.78)$$

Podstawiając  $L = \sqrt{\mu a}$  oraz  $G = \sqrt{\mu p} = \sqrt{\mu a (1 - e^2)}$ , możemy dla dowolnej funkcji  $\phi$  zmiennych  $a, e, I$  napisać

$$\frac{\partial \phi(a, e, I)}{\partial L} = 2 \sqrt{\frac{a}{\mu}} \frac{\partial \phi}{\partial a} + \frac{1 - e^2}{e \sqrt{\mu a}} \frac{\partial \phi}{\partial e},$$

$$\begin{aligned}\frac{\partial\phi(a, e, I)}{\partial G} &= -\frac{\sqrt{1-e^2}}{e\sqrt{\mu a}} \frac{\partial\phi}{\partial e} + \frac{c}{s\sqrt{\mu p}} \frac{\partial\phi}{\partial I}, \\ \frac{\partial\phi(a, e, I)}{\partial H} &= -\frac{1}{s\sqrt{\mu p}} \frac{\partial\phi}{\partial I}.\end{aligned}\quad (3.79)$$

Dysponując wszystkimi niezbędnymi pochodnymi cząstkowymi możemy przystąpić do formułowania równań ruchu dla zmiennych keplerowskich.

### 3.4.2 Równania dla zmiennych keplerowskich

Zacznijmy od równań dla kątów czyli  $M, \omega, \Omega$ . Co prawda nachylenie  $I$  jest też kątem ale widzimy, że spokrewnione z pędami  $G$  i  $H$ , dynamicznie należy ono do tej samej rodziny co półoś wielka i mimośród. W świetle wzorów (3.75) – (3.79), dla anomalii średniej mamy

$$\begin{aligned}\dot{M} &= \{M, \mathcal{H}\} = \frac{\partial M}{\partial \ell} \frac{\partial \mathcal{H}}{\partial L} = 2\sqrt{\frac{a}{\mu}} \frac{\partial \mathcal{H}}{\partial a} + \frac{1-e^2}{e\sqrt{\mu a}} \frac{\partial \mathcal{H}}{\partial e} = \\ &= \sqrt{\frac{\mu}{a^3}} + 2\sqrt{\frac{a}{\mu}} \frac{\partial V}{\partial a} + \frac{1-e^2}{e\sqrt{\mu a}} \frac{\partial V}{\partial e}.\end{aligned}$$

Dla argumentu perycentrum

$$\dot{\omega} = \{\omega, \mathcal{H}\} = \frac{\partial \omega}{\partial g} \frac{\partial \mathcal{H}}{\partial G} = -\frac{\sqrt{1-e^2}}{e\sqrt{\mu a}} \frac{\partial V}{\partial e} + \frac{c}{s\sqrt{\mu p}} \frac{\partial V}{\partial I},$$

natomiast dla długości węzła wstępującego

$$\dot{\Omega} = \{\Omega, \mathcal{H}\} = \frac{\partial \Omega}{\partial h} \frac{\partial \mathcal{H}}{\partial H} = -\frac{1}{s\sqrt{\mu p}} \frac{\partial V}{\partial I}.$$

Z kolei dla półosi wielkiej, będącej jedynie funkcją  $L$ ,

$$\dot{a} = \{a, \mathcal{H}\} = -\frac{\partial a}{\partial L} \frac{\partial \mathcal{H}}{\partial \ell} = -2\sqrt{\frac{a}{\mu}} \frac{\partial V}{\partial M}.$$

Mimośród jest funkcją  $L$  i  $G$ , a zatem

$$\dot{e} = \{e, \mathcal{H}\} = -\frac{\partial e}{\partial L} \frac{\partial \mathcal{H}}{\partial \ell} - \frac{\partial e}{\partial G} \frac{\partial \mathcal{H}}{\partial g} = -\frac{1-e^2}{e\sqrt{\mu a}} \frac{\partial V}{\partial M} + \frac{\sqrt{1-e^2}}{e\sqrt{\mu a}} \frac{\partial V}{\partial \omega}.$$

I wreszcie, dla nachylenia  $I$ , zależnego od  $G$  i  $H$ , otrzymujemy

$$\dot{I} = \{I, \mathcal{H}\} = -\frac{\partial I}{\partial G} \frac{\partial \mathcal{H}}{\partial g} - \frac{\partial I}{\partial H} \frac{\partial \mathcal{H}}{\partial h} = -\frac{c}{s\sqrt{\mu p}} \frac{\partial V}{\partial \omega} + \frac{1}{s\sqrt{\mu p}} \frac{\partial V}{\partial \Omega}.$$

Zbierzmy teraz sześć wyprowadzonych powyżej równań planetarnych Lagrange'a. Wprowadźmy oprócz  $c = \cos I$  i  $s = \sin I$  symbole

$$n = \sqrt{\mu/a^3}, \quad \eta = \sqrt{1 - e^2}.$$

Wielkość  $n$  jest oczywiście tylko wygodnym skrótem i nie oznacza, że w zagadnieniu zaburzonym obowiązuje nadal III prawo Keplera. Równania Lagrange'a możemy zapisać w uproszczonej postaci

$$\begin{aligned} \dot{a} &= -\frac{2}{na} \frac{\partial V}{\partial M}, \\ \dot{e} &= -\frac{\eta}{na^2 e} \left( \eta \frac{\partial V}{\partial M} - \frac{\partial V}{\partial \omega} \right), \\ \dot{I} &= -\frac{1}{na^2 s \eta} \left( c \frac{\partial V}{\partial \omega} - \frac{\partial V}{\partial \Omega} \right), \\ \dot{M} &= n + \frac{2}{na} \frac{\partial V}{\partial a} + \frac{\eta^2}{na^2 e} \frac{\partial V}{\partial e}, \\ \dot{\omega} &= -\frac{1}{na^2} \left( \frac{\eta}{e} \frac{\partial V}{\partial e} - \frac{c}{s \eta} \frac{\partial V}{\partial I} \right), \\ \dot{\Omega} &= -\frac{1}{na^2 \eta s} \frac{\partial V}{\partial I}. \end{aligned} \tag{3.80}$$

Spróbujmy zestawić równania Lagrange'a z równaniami Gaussa (1.70)–(1.75). Równania Lagrange'a stosują się do węższej klasy sił, ale za to mają zasadniczo prostszą postać i pozwalają na wysuwanie ogólniejszych wniosków z samego wglądu w strukturę równań ruchu. O ruchu decyduje tutaj pojedyncza funkcja skalarna  $V$ . Warto pamiętać, że w niektórych źródłach równania (3.80) występują z funkcją sił  $U = -V$  zamiast potencjału i wtedy różnią się znakami od podanych tutaj.

W porównaniu z równaniami kanonicznymi w zmiennych Delaunaya (3.73), niekanoniczne równania Lagrange'a są mniej proste, ale z drugiej strony łatwiej jest wyrażać potencjał  $V$  przy użyciu zmiennych keplerowskich niż przy użyciu zmiennych Delaunaya. Ze względu na powiązania zmiennych  $M, \omega, \Omega$  z kątami  $\ell, g, h$ , zaś zmiennych  $a, e, I$  z pędami  $L, G, H$ , równania Lagrange'a do pewnego stopnia dziedziczą resztki struktury równań (3.73). Zauważmy, że istotnie równania (3.80) rozpadają się na dwie podgrupy: w równaniach dla zmiennych  $a, e, I$  występują jedynie pochodne potencjału względem  $M, \omega, \Omega$  i *vice versa*. Gdyby więc potencjał zaburzający  $V$  nie zależał jawnie od czasu ani od kątów  $M, \omega$  i  $\Omega$ , to wtedy elementy  $a, e, I$  byłyby stałe, co z kolei pociągałoby za sobą stałość prędkości kątowych  $\dot{M}, \dot{\omega}$

i  $\dot{\Omega}$ . W słabszych przypadkach szczególnych mielibyśmy: dla  $\frac{\partial V}{\partial t} = \frac{\partial V}{\partial M} = 0$  – stałą półoś wielką  $a$ , zaś dla  $\frac{\partial V}{\partial t} = \frac{\partial V}{\partial \Omega} = \frac{\partial V}{\partial \omega} = 0$  – stałe nachylenie  $I$ .

Na zakończenie przypomnijmy, że do równań Lagrange’a (3.80), podobnie jak do równań kanonicznych (3.73) i równań Gaussa (1.70)–(1.75), odnoszą się tradycyjne zastrzeżenia o możliwości pojawienia się osobliwości dla małych nachyleń, małych mimośrodów, dla orbit zdegenerowanych i dla orbit prawie parabolicznych.

# Spis treści

|          |  |           |
|----------|--|-----------|
| <b>1</b> | <b>Równania ruchu w formalizmie newtonowskim</b>                             | <b>1</b>  |
| 1.1      | Formalizm newtonowski . . . . .  | 1         |
| 1.2      | Wahadło matematyczne, oscylator harmoniczny i oscylator Duffinga . . . . .   | 4         |
| 1.3      | Metoda uzmienniania stałych . . . . .  | 7         |
| 1.4      | Uzmiennianie stałych dla oscylatora Duffinga . . . . .                       | 9         |
| 1.5      | Elementy oskulacyjne i równania Gaussa . . . . .                             | 13        |
| 1.5.1    | Uzmiennione całki ruchu . . . . .  | 14        |
| 1.5.2    | Równanie dla pólوسی wielkiej . . . . .                                       | 15        |
| 1.5.3    | Równanie dla mimośrodu . . . . .   | 16        |
| 1.5.4    | Kąty Eulera a prędkość kątowna . . . . .                                     | 16        |
| 1.5.5    | Równania dla oskulacyjnej długości węzła wstępującego i nachylenia . . . . . | 18        |
| 1.5.6    | Równanie dla oskulacyjnego argumentu perycentrum . . . . .                   | 20        |
| 1.5.7    | Równania dla zmian anomalii . . . . .  | 21        |
| 1.5.8    | Składowe siły zaburzającej . . . . .   | 23        |
| 1.5.9    | Ostateczna postać wzorów Gaussa . . . . .                                    | 26        |
| 1.5.10   | Elementy nieosobliwe . . . . .   | 28        |
| 1.6      | Przykładowe zastosowania równań Gaussa . . . . .                             | 30        |
| 1.6.1    | Wpływ oporu ośrodka na elementy oskulacyjne. . . . .                         | 30        |
| 1.6.2    | Manewry orbitalne (zmiana nachylenia) . . . . .                              | 32        |
| <b>2</b> | <b>Formalizm kanoniczny</b>  | <b>34</b> |
| 2.1      | Macierze i formy symplektyczne . . . . .                                     | 34        |
| 2.1.1    | Standardowa macierz symplektyczna i forma symplektyczna . . . . .            | 34        |
| 2.1.2    | Definicja macierzy symplektycznych . . . . .                                 | 35        |
| 2.1.3    | Grupa symplektyczna . . . . .  | 36        |
| 2.2      | Pfaffian i równania Pfaffa . . . . .   | 39        |

|          |   |           |
|----------|---|-----------|
| 2.3      | Równania kanoniczne jako przypadek szczególny równań Pfaffa             | 41        |
| 2.4      | Kanoniczne nawiasy Poissona   | 45        |
| 2.5      | Równania kanoniczne dla układu mechanicznego                            | 47        |
| 2.5.1    | Układ $K$ ciał  | 47        |
| 2.5.2    | Zagadnienie dwóch ciał  | 49        |
| 2.5.3    | Ruch pojedynczej cząstki w polu sił                                     | 49        |
| <b>3</b> | <b>Transformacje kanoniczne</b>   | <b>51</b> |
| 3.1      | Podstawy  | 51        |
| 3.1.1    | Definicja transformacji kanonicznej                                     | 51        |
| 3.1.2    | Warunek dostateczny kanoniczności transformacji                         | 52        |
| 3.1.3    | Nawiasy Poissona a transformacja kanoniczna                             | 55        |
| 3.2      | Transformacje Mathieu   | 58        |
| 3.2.1    | Definicja   | 58        |
| 3.2.2    | Liniowa transformacja kanoniczna  | 59        |
| 3.2.3    | Ruch w układzie jednostajnie obracających się osi                       | 61        |
| 3.2.4    | Zagadnienie względne dwóch ciał   | 64        |
| 3.2.5    | Zmienne Hilla-Whittakera  | 67        |
| 3.3      | Transformacje zadane funkcją tworzącą                                   | 71        |
| 3.3.1    | Funkcja tworząca zmiennych mieszanych                                   | 71        |
| 3.3.2    | Transformacja tożsamościowa   | 72        |
| 3.3.3    | Zmienne ką-t-działanie  | 73        |
| 3.3.4    | Zmienne Delaunaya   | 76        |
| 3.4      | Równania planetarne Lagrange'a  | 81        |
| 3.4.1    | Pochodne cząstkowe zmiennych keplerowskich względem zmiennych Delaunaya | 82        |
| 3.4.2    | Równania dla zmiennych keplerowskich                                    | 83        |